U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

MSc

MSc

2.º
CICLO

FCUP
UV
EM
2018

U.PORTO

Establishing the Mechanism of Glutaredoxin

Sodiq Odewale Waheed

FC

U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

VNIVERSITAT DE VALÈNCIA

European Commission
ERASMUS
MUNDUS

# Establishing the Mechanism of Glutaredoxin

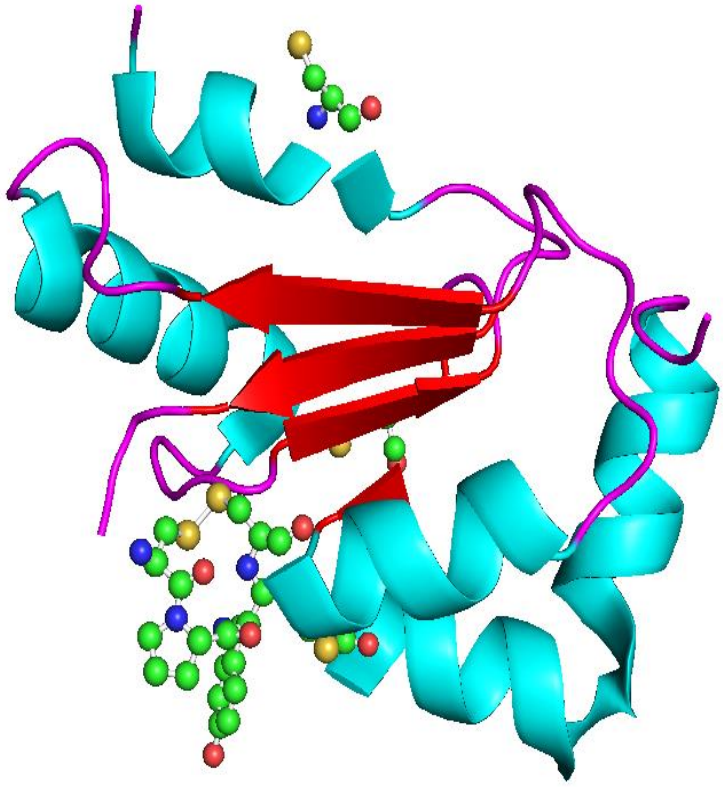
Sodiq Odewale Waheed

Dissertação de Mestrado apresentada à
Faculdade de Ciências da Universidade do Porto
Química

2018

MSc

MSc



# Establishing the Mechanism of Glutaredoxin


Sodiq Odewale Waheed

Mestrado em Química

Departamento de Química e Bioquímica

2018

## Orientadora

Maria João Ramos, Professora Catedrática, Faculdade de Ciências da Universidade do Porto

## Coorientadores

Sérgio F. Sousa, Investigador, Faculdade de Ciências da Universidade do Porto

Iñaki Tuñón, Full Professor, Faculty of Chemistry, University of Valencia, Spain




Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____

# ACKNOWLEDGEMENTS

First and foremost, I am heartily grateful to the Erasmus Mundus Programme of the European Master in Theoretical Chemistry and Computational Modelling (EMTCCM) with the support of the European Union for the Erasmus Mundus grant they awarded me for this Master Programme.

I am extremely indebted to my supervisor – Professor Maria João Ramos who despite her hectic schedule gave all the needed guidance and help necessary for the successful completion of this thesis and the Master programme in general. My earnest desire is that God will continue to keep you at his right hand.

My profound gratitude also goes to Professor Pedro A. Fernandes, Dr. Sérgio F. Sousa and Dr. Rui P.P. Neves for their innovative ideas and support from the beginning to the end of this research. May the good Lord bless you all.

I would like to render my sincere appreciations to all the Professors in the Department of Chemistry and Biochemistry, FCUP and all the Professors during the Intensive Course part of the Master at the University of Valencia, Spain, as well as during the TCCM winter School at the University of Toulouse III - Paul Sabatier, France, who have in one way or the other imparted in me the skills needed for a competitive edge in life.

I wish to also express my gratitude to Professor Iñaki Tuñón's research group at the Faculty of Chemistry, University of Valencia, Spain for the privilege that was given to me to carry out part of my research project in their group. I am highly thankful to Professor Iñaki Tuñón, Kirill Zinovjev and Carlos Ramos-Guzmán.

To my amiable and loving family miles away, thanks so much for your interminable love, support, prayers, and encouragements.

Finally, I thank God for His unfeigned love and unprecedented faithfulness during this Master programme.




# ABSTRACT

Glutaredoxin (Grx) is a redox enzyme and it belongs to thioredoxin-family with Cys-Pro-Tyr-Cys active site motif. This enzyme catalyzes the reduction of glutathione disulfide (GSSG) to glutathione (GSH) through the dithiol mechanism that includes a Grx-S-S-G mixed disulfide intermediate between the enzyme (Grx) and the substrate (GSSG). Glutaredoxin functions as both electron donors and as regulators of cellular function in response to oxidative stress. Additionally, it functions in sulfur assimilation, dehydroascorbate reduction, and the regulation of cellular differentiation, transcription and apoptosis. Grx and other redox enzymes play a very crucial role in the protection of dopaminergic neurons. The loss of dopaminergic neurons results in Parkinson's disease; hence, the study of this enzyme becomes important.

The work presented in this thesis contains the outcomes of the studies of the catalytic mechanism of the reduction of glutathione disulfide (GSSG) by glutaredoxin (Grx) using the ONIOM extrapolative QM/MM scheme as implemented in Gaussian09 package. The whole thesis contains four chapters. Chapter one presents an introductory information about the enzyme, its characteristic reaction; thiol-disulfide exchange, and the catalytic mechanism of the enzyme. Chapter two dwells on discussing the computational methods as applied to biological systems. This chapter contains four sections: (i) Quantum Mechanics (QM) Methods, (ii) Molecular Mechanics (MM) Methods, (iii) Hybrid Methods with emphasis on QM/MM Methods, and (iv) Computational Enzymatic Catalysis. Chapter three describes the studies that were carried out with the results obtained for the two mechanistic steps of the reaction as well as the conclusions drawn from each step. Lastly, chapter four outlines the general conclusions from the whole work.

**Keywords:** Glutaredoxin, quantum mechanics, thioredoxin-family, glutathione (GSH), molecular mechanics, enzymatic catalysis, oxidized glutathione (GSSG), thiol-disulfide exchange, ONIOM QM/MM, dithiol mechanism.

# RESUMO

A glutaredoxina (Grx) é uma enzima redox e pertence à família das tiorredoxinas com o motivo activo Cys-Pro-Tyr-Cys. Esta enzima catalisa a redução de dissulfeto de glutationa (GSSG) a glutationa (GSH) através do mecanismo de ditiol que inclui um intermediário dissulfeto de Grx-S-S-G entre a enzima (Grx) e o substrato (GSSG). A glutaredoxina funciona como dadora de eletrões e como reguladora da função celular em resposta ao stressoxidativo. Além disso, funciona na assimilação de enxofre, na redução do desidroascorbato e na regulação da diferenciação celular, transcrição e apoptose. O grx e outras enzimas redox desempenham um papel crucial na proteção dos neurônios dopaminérgicos. A perda de neurônios dopaminérgicos resulta na doença de Parkinson, tornando o estudo desta enzima importante.

O trabalho apresentado nesta tese contém os resultados dos estudos do mecanismo catalítico da redução da glutationa (GSSG), por glutarredoxina (Grx) usando o esquema ONIOM extrapolativo QM / MM como implementado no pacote Gaussian09. A tese contém no seu todo quatro capítulos.O capítulo um apresenta informação introdutória sobre a enzimae a sua reação característica; a permuta de tiol-dissulfureto, e o mecanismo catalítico da enzima. O capítulo dois inclui a discussão dos métodos computacionais e a sua aplicação a sistemas biológicos. Este capítulo contém quatro secções: (i) Métodos de Mecânica Quântica (QM), (ii) Métodos de Mecânica Molecular (MM), (iii) Métodos Híbridos com ênfase nos Métodos QM / MM, e (iv) Catálise Computacional Enzimática. O capítulotrês descreve os estudos que foram efetuados com os resultados obtidos para os dois passos mecanísticos da reação, bem como as conclusões retiradas a partir de cada etapa. Por fim, o capítulo quatro descreve as conclusões gerais de todo o trabalho.

**Palavras-chave:** Glutaredoxina, mecânica quântica, tiorredoxina-família, glutationa (GSH), mecânica molecular, catálise enzimática, glutationa oxidada (GSSG), permuta de tiol-dissulfureto, ONIOM QM / MM, mecanismo ditiol.

# TABLE OF CONTENTS

# INDEX OF FIGURES

# INDEX OF TABLE

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AM1 | Austin Model 1 |
| AMBER | Assisted Model Building with Energy Refinement |
| CHARMM | Chemistry at HARvard Macromolecular Mechanics |
| CPMD | Car-Parrinello Molecular Dynamics |
| CPU | Central Processing Unit |
| DDT | Dithiothreitol |
| DFT | Density Functional Theory |
| DsbA | Disulfide bond formation A |
| EVB | Empirical Valence Bond |
| GAFF | General Amber Force Field |
| GGA | Generalized Gradient Approximation |
| GHO | Generalized Hybrid Orbital |
| GR | Glutathione Reductase |
| Grx | Glutaredoxin |
| GSH | Glutathione |
| GSSG | Glutathione Disulfide |
| GTO | Gaussian-type orbital |
| HED | Hydroxy ethyl disulfide |
| HF | Hartree Fock |
| hGrx | Human Glutaredoxin |
| HIV | Human Immunodeficiency Virus |
| IMOMM | Integrated Molecular Orbital and Molecular Mechanics |
| IMOMO | Integrated Molecular Orbital and Molecular Orbital |
| INT | Reaction Intermediates |
| kDa | kilo Dalton |
| LDA | Local Density Approximation |
| LSCF | Localized Self Consistent Field |
| MD | Molecular Dynamics |
| MM | Molecular Mechanics |

| | |
|---|---|
| MNDO | Modified Neglect of Diatomic Differential Overlap |
| NADPH | Nicotinamide adenine dinucleotide phosphate |
| NMR | Nuclear Magnetic Resonance |
| NPT | Isobaric-isothermal ensemble |
| NVT | Canonical ensemble |
| ONIOM | Our own N-layered Integrated Molecular Orbital and Molecular Mechanics |
| OPLS | Optimized Potential Liquid Simulation |
| OPLS-AA | OPLS- All Atoms |
| OPLS-UA | OPLS-United Atoms |
| P | Product |
| PD | Parkinson's Disease |
| PDB | Protein Data Bank |
| PDI | Protein Disulfide Isomerase |
| PES | Potential Energy Surface |
| PM3 | Parametric Model number 3 |
| PM6 | Parametric Model number 6 |
| QM | Quantum Mechanics |
| QM/MM | Quantum Mechanics/Molecular Mechanics |
| R | Reactant |
| RNR | Ribonucleotide Reductase |
| SCC-DFTB | Self consistent charge Density Functional tight binding |
| SP | Single Point |
| TS | Transition State |
| STO | Slater-type orbital |
| TIP3P | Transferable Intermolecular Potential 3-Point |
| Trx | Thioredoxin |
| ZPE | Zero-point Energy |

# AMINO ACIDS ABBREVIATIONS

| | |
|---|---|
| Ala | Alanine |
| Arg | Arginine |
| Asn | Asparagine |
| Asp | Aspartate |
| Cys | Cysteine |
| Gln | Glutamine |
| Glu | Glutamate |
| Gly | Glycine |
| His | Histidine |
| Ile | Isoleucine |
| Leu | Leucine |
| Lys | Lysine |
| Met | Methionine |
| Phe | Phenylalanine |
| Pro | Proline |
| Ser | Serine |
| Thr | Threonine |
| Trp | Tryptophan |
| Tyr | Tyrosine |
| Val | Valine |

# Chapter One – Biological Problem

## 1.0 Introduction

The chemistry of cysteine is the key to the enzymatic mechanism of the thiol-disulfide oxidoreductases of the thioredoxin subfamily, such as thioredoxins (Trxs), glutaredoxins (Grxs), protein disulfide isomerase (PDI), and DsbA [1]. These enzymes are ubiquitous and share similar architecture, known as the thioredoxin fold [2-3]. Trx, Grx, PDI, and DsbA active sites contain a -Cys-X-X-Cys- motif, where the N-terminal cysteine thiol has a $pK_a$ that is significantly lower than that of a free cysteine [4-11]. Conversely, the $pK_a$ of the C-terminal active site cysteine thiol is usually more basic than that of a free cysteine [1]. The understanding of the thiol-disulfide reaction is important because it is involved in a number of biological processes ranging from protein folding and stability to regulation of gene expression and catalytic activity [5-7].

Depending on the protein under study, measured values of the $pK_a$ of the N-terminal active site cysteine range from ~7.1 in Trx [7,12,13] to ~3.2 in DsbA [9,14]. The thiol sulfur reactivity is highly dependent on the $pK_a$ of the thiol, as a lower $pK_a$ affects both nucleophilicity and leaving group ability of the associated thiolate [15,16]. This is well illustrated in DsbA and its mutants, where a lower $pK_a$ of the N-terminal cysteine of the -Cys-X-X-Cys- motif is associated with a more oxidative protein [14,17,18]. Although, other factors influence the redox properties of the -Cys-X-X-Cys- motif, the $pK_a$ of the N-terminal cysteine clearly contributes to these properties [19]. In addition, the nature of the -X-X- residues also has crucial effects on the N-terminal cysteine thiol $pK_a$ and is mostly understood in terms of the number of hydrogen bonds that the -X-X- residues can provide to stabilize the thiolate charge [1]. The replacement of a Tyr in the -Cys-Pro-Tyr-Cys- motif of Grx (or a His in the -Cys-Pro-His-Cys- motif of DsbA) by a Pro in its Trx counterpart (Cys-Gly-Pro-Cys-) allows the backbone of Grx (DsbA) to donate one more hydrogen bond to the thiolate sulfur than in Trx, which lowers the $pK_a$ of the N-terminal active site cysteine in Grx (DsbA) more than in Trx [20,21]. The His of the -Cys-Pro-His-Cys- motif in DsbA can also form a hydrogen bond with the thiolate sulfur [22], which may also stabilize the thiolate anion in DsbA more than in Grx [20]. The thiol of the C-terminal cysteine is another group that can form a hydrogen bond with the thiolate [13]. The reduced Trx [23] and the reduced DsbA [22] X-ray crystal structures are compatible with such a hydrogen bond. The

direct evidence that supports this hydrogen bond is provided by the NMR experiments [10] and molecular dynamics (MD) simulations [20] using glutaredoxin 3 (Grx3) from *Escherichia coli*.

# 1.1 Glutaredoxins

The Glutaredoxin system was initially described by Holmgren in 1976 as an alternate electron donor to ribonucleotide reductase (RNR) in *Escherichia coli* mutants lacking thioredoxin [24]. A few years later, Grx was shown to also have GSH-disulfide transhydrogenase activity; thus, functioning as general thiol-disulfide oxidoreductases [25]. In the end, it became apparent that thioltransferases and glutaredoxins purified from different organisms were highly similar and were actually the same enzymes just named differently [5,11].

Glutaredoxins are small proteins, usually around 9-15 kDa, that exist in a large number of isoforms in basically all glutathione (GSH) containing life forms [26]. According to their active site motif, Grxs can be divided in two major categories: the dithiol Grxs with active site motif Cys-X-X-Cys and the monothiol Grxs with active site motif Cys-X-X-Ser [27,28]. The latter is subdivided into single- and multi-domain monothiol Grxs [27,28]. Dithiol and single-domain monothiol Grxs are present in all forms of life such as bacteria, plants, and vertebrates, while multi-domain Grxs were until now not yet identified in bacteria [29]. In addition, another division of the Grx family is based on a cysteine residue at the second position of the active site motif (Cys-Cys-X-Cys/Ser). These Grxs family, called CC-type Grxs, are exclusively found in land plants, where they control plant physiology, like flower development [30]. Dithiol Grxs are thiol-disulfide oxidoreductases that have high specificity for the glutathione moiety [31]. These Grxs are kept reduced by GSH, and the resulting oxidized form of GSH, glutathione disulfide (GSSG) is reduced by glutathione reductase (GR) with electrons from NADPH [31,32]. The classical dithiol Grxs efficiently reduce some protein disulfides like that in *Escherichia coli* ribonucleotide reductase with a dithiol mechanism [26]. However, the major action of Grxs appears to be to reduce protein-GSH mixed-disulfides using a mechanism that requires only the N-terminal active site residue, that is, a monothiol mechanism [33].

Grxs have been suggested to be involved in numerous functions, both as electron donors and as regulators of cellular function in response to oxidative stress, as demonstrated by the ability of yeast glutaredoxins to directly reduce hydroperoxides with catalytic rate comparable to glutathione peroxidase [34], and the sensitivity of the yeast mutants lacking Grx to oxidative stress [35]. Additionally, they function in sulfur assimilation [36,37],

dehydroascorbate reduction [38], and the regulation of cellular differentiation [39], transcription [40-42], and apoptosis [43-45]. The human Grx2 can also catalyze the reduction of hydroperoxides *in vitro*, but a reasonable high concentration of the enzyme is needed [46]. Furthermore, human Grx1 has been shown to be able to scavenge glutathione-thiyl radicals [47]. Grxs also mediate the recovery from oxidative damage by reducing protein-GSH mixed disulfides and restoring the protein thiol homeostasis in brain mitochondria [48], or human lens epithelial cells [49].

A study demonstrating the critical role of the glutaredoxin system in protecting macrophages from cell death induced by oxidized low density lipoprotein, was used to show how Grxs protect cells against apoptosis [50]. In addition, glutaredoxins have been detected within HIV-1 and have been proven to regulate the activity of glutathionylated HIV-1 protease *in vitro* [51]. Monothiol Grxs from *E. coli*, yeast and vertebrates including humans are crucially involved in iron-sulfur cluster biosynthesis and regulation of iron homeostasis [52].

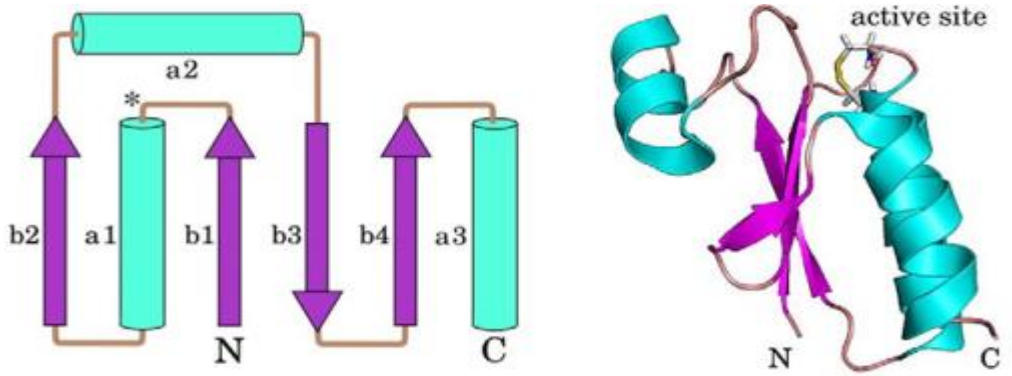
The study of this enzyme becomes important, as it has been shown with other redox enzymes, to play a very crucial role in the protection of dopaminergic neurons, thus contributing to maintain neuronal cell viability. The loss of dopaminergic neurons results in the noxious disease called Parkinson's disease (PD).

# 1.2 Structure of Glutaredoxins

Glutaredoxins have been extensively studied using both NMR spectroscopy and X-ray crystallography. There are over one hundred (100) structures of Grxs from different organisms, like *Homo sapiens*, *E. coli*, *Saccharomyces cerevisiae*, etc. in the protein database. Structurally, Grxs belong to the Trx fold family of proteins, and represent only a substructure or a domain in the other members of the family [2,3]. In bacterial glutaredoxins, the structure contains four stranded β-sheets that are surrounded by three α-helices (Fig. 1A), while other Grx types from yeast, human etc. do consist of four stranded β-sheets surrounded by five α-helices (Fig. 1B). The central β-sheet is composed of three antiparallel strands (b1,b3,b4) with b2 parallel to the adjacent b1 (Fig. 1). Furthermore, all the Trx oxidoreductase family of proteins shares similar Cys-X-X-Cys active site motifs that are located on the loop connecting the first β-sheet (b1) and the first α-helix (a1). The N-terminal cysteine residue in the Grx active site is similar to Trxs, surface-exposed with low pKa, below that of the free cysteine, while the C-terminal cysteine is buried in the molecule and has much higher pKa value [28].

Human glutaredoxin consists of a four-stranded mixed β-sheet composed of residues 15 to 19, 43 to 47, 72 to 75 and 78 to 81, and five α-helices composed of residues 4 to 9, 24 to 34, 54 to 65, 83 to 91 and 94 to 100 [53] in contrast to the three α-helices that are observed in *E. coli* Grx-1. The two additional helices in hGrx relative to its *E. coli* counterpart are located at the N and C termini of the protein [53]. The loop connecting α-helix 2 to β-strand 2 (α1 to β2 in *E. coli*) is longer in the human protein. It was reported that a hydrogen bond interaction between the NH of Lys39 and the CO of Ala104 serves to hold α-helix 5 in contact with the extended loop between α-helix 2 to β-strand 2 [53]; thus, enhances the interaction between the extended loop and α-helix 5 [53]. The helix containing the active site is one turn shorter in the human protein [53]. Also, α-helix 3 in hGrx is one turn longer and positioned differently than the counterpart *E. coli* Grx-1. This has been reported to result in the narrowing of the groove into which the RNR substrate has been shown to bind in *E. coli* Grx-1 [54]. The β-sheet forms the central core of the protein with helices 1 and 3 located on one side of the sheet and helices 2, 4, and 5 located on the other side [53]. Human Grx (hGrx) needs to be structurally characterized because it is a basic protein and longer than its *E. coli* counterpart. Furthermore, human Grxs have three additional cysteine residues (Cys8, Cys79 and Cys 83) in addition to those that are present in the active site [53]. These additional cysteines are proposed to play a regulatory role because their oxidation leads to inactivation in the case of the studied calf thymus Grx [55]. The exact role of these additional cysteine residues is yet to be established [53]. However, possible functional importance was suggested for them based on their spatial location and local environment [53]. Cys8 which is located in the N-terminal helix and solvent-exposed, decreases the lifetime of hGrx samples considerably with substantial aggregation of the protein if it is not maintained in a fully reduced state [53]. The second additional cysteine residue (Cys79) is also solvent-exposed and is located in β-strand 4, providing a site for oligomerization. Oligomerization of hGrx with disulfide cross-linking helps in the regulation of hGrx [53]. The last cysteine residue, Cys83 is located at the N-terminal of α-helix 4 and is solvent-exposed too [53]. Cys83 is located in a region of positive electrostatic potential, which is indicative that it will have a low pKa value and that likely will be active in redox reactions. The thiol would play a key role in the regulation of the protein or in one or more known activities of the protein [53].

**A**




**B**

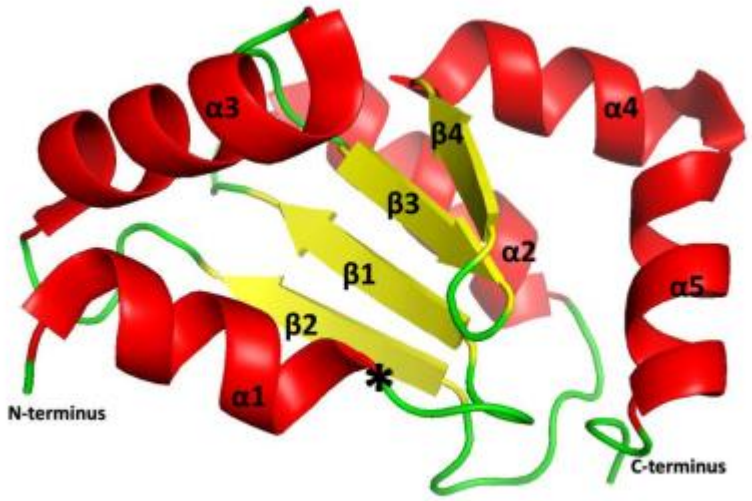



Fig.1: Glutaredoxins structures from bacteria (A) [28] and other from human (B) [161]. The notation * in A and B represent the point where the Cys-Pro-Tyr-Cys active site motif is located in the enzyme.

# 1.3 Catalytic Mechanism of Glutaredoxins

The catalytic mechanism of Grxs has been reported to follow the thiol-disulfide exchange reaction, characteristic of the thioredoxin oxidoreductases family [10,26,28-33,56-61,64].

## 1.3.1 Thiol-Disulfide Exchange Reaction

The thiol-disulfide exchange reaction is a biological fundamental process [56]. This reaction is reversible and reaches an equilibrium based on the initial concentrations of the reactants and products, governed by the redox potentials of the thiol-disulfide couples involved [57]. Thiol-disulfide reaction plays a very crucial role in protein folding and stability [56], and the folding rate is usually limited by this reaction for proteins containing disulfide bonds [56]. The reaction can be represented by the set of chemical equations below [56]:

$$R\text{-}SH + H_2O \rightleftharpoons R\text{-}S^- + H_3O^+ \quad \textbf{(1)}$$

$$R\text{-}S^- + R'\text{-}SS\text{-}R' \rightleftharpoons R\text{-}SS\text{-}R' + R'\text{-}S^- \quad \textbf{(2)}$$

$$R\text{-}SS\text{-}R' + R\text{-}S^- \rightleftharpoons R\text{-}SS\text{-}R + R'\text{-}S^- \quad \textbf{(3)}$$

$$R'\text{-}S^- + H_3O^+ \rightleftharpoons R'\text{-}SH + H_2O \quad \textbf{(4)}$$

The overall reaction is:

$$2\ R\text{-}SH + R'\text{-}SS\text{-}R' \rightleftharpoons R\text{-}SS\text{-}R + 2\ R'\text{-}SH \quad \textbf{(5)}$$

In the above overall reaction, a reduced thiol (R-SH) is exchanged with a disulfide (R´-SS-R´), leading to formation of a new disulfide (R-SS-R) and a new thiol (R´-SH). Although, the reaction is a redox reaction, it occurs as a nucleophilic displacement, with the thiol nucleophile attacking the electrophile disulfide. The rate of this reaction depends on the nucleophilicity of the thiol reactant, the reactivity of the central sulfur atom being attacked and the stability of the leaving group thiol product [56,57].

The overall reaction of Grx mediated by thiol-disulfide exchange follows a ping-pong mechanism [58]. Two different interconnected mechanisms have been described, the dithiol and the monothiol mechanism. In the dithiol mechanism, both active site cysteine residues of the Grxs are used to catalyze the reversible reduction of the disulfide while the monothiol mechanism utilizes only the N-terminal active site cysteine and reduces mixed disulfides formed between GSH and proteins or other small thiol compounds. Monothiol Grxs are not able to

perform the reduction of disulfide because of the lack of the C-terminal active site cysteine (buried cysteine) [29].

The thiol-disulfide exchange activity of Grxs can be determined by a limited set of assays. The experimentalists use these set of assays to distinguish between the monothiol and dithiol mechanisms in the laboratory [29]. The Grx deglutathionylation activity is usually measured using an artificial non-specific substrate HED (β-hydroxyethyl disulfide) in a spectrophotometric coupled assay measuring the consumption of NADPH by GR as described by Holmgren [25]. In addition, other substrates like cysteine-glutathione mixed disulfides or radiolabelled glutathionylated proteins are used for measuring the activity either coupled with NADPH or by radioactive determinations [17,58]. The disulfide reduction activity is measured using a ribonucleotide reductase (RNR) assay [25]. This assay measures the formation of $^{3}$H-dCDP from $^{3}$H-CDP by RNR using electrons provided by GSH, GR and NADPH or with all the three replaced with dithiothreitol (DDT) [25].

## 1.3.2 Thiol-Disulfide Exchange Catalytic Mechanism of Glutaredoxin

Glutaredoxin like other thioredoxin-like enzyme catalyzes the reduction of oxidized GSH (GSSG) to a reduced GSH (GSH) [59] via the dithiol mechanism that includes a mixed disulfide between Grx and substrate [60,61]. This mechanism is often regarded as ping-pong mechanism [58,62,63]. The catalytic mechanism scheme is shown in Fig. 2.

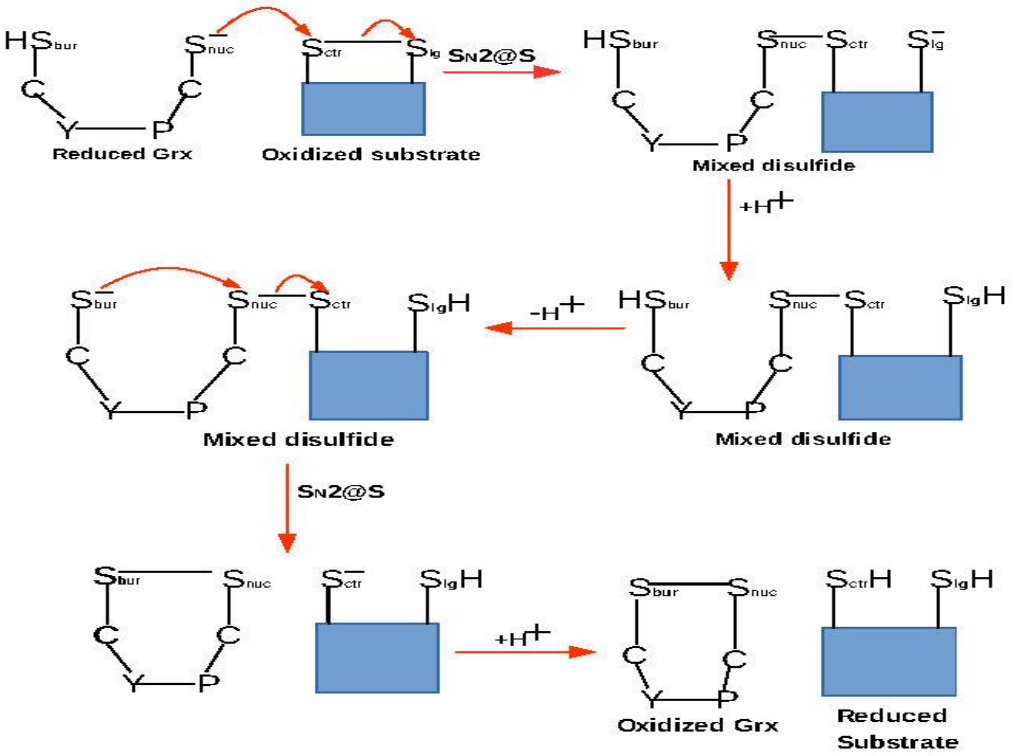



Fig. 2: Catalytic Mechanism Scheme for the reduction of disulfide with Grx [26,28,29,33,57,63,64].

As shown above, the mechanism involves two basic steps. In the first, the nucleophilic sulfur atom of the N-terminal active site cysteine is used to attack the central sulfur atom of the substrate disulfide bond. The cysteine has low pKa value, solvent accessible, and it is primed to attack the disulfide bond of the oxidized substrate [57], while the C-terminal cysteine is inaccessible and buried [29]. Due to the low pKa value of the N-terminal active site cysteine when compared with the free cysteine, it is present as thiolate at physiological pH [11], and a larger fraction of the anionic thiolate would be available to react. Most reactive thiols are those with a pKa close to the pH of the solution [56]. The thiol-disulfide exchange reaction rate constant is dependent on the nucleophilicity of the attacking reactant and the stability of the leaving group. Thiolates have been proven to be better nucleophiles than thiols, and this explains the catalytic efficiency of Trx family in thiol-disulfide exchange reaction at physiological pH [29]. The second step involves the use of the C-terminal cysteine (buried cysteine) to reduce

the mixed disulfide intermediate that is formed in the first step. This step releases a reduced substrate and an oxidized protein with disulfide bond between the two active site cysteines [29].

The abstraction or deprotonation of the hydrogen from Cys-$S_{nuc}$H in the first step may be due to the fact that the nucleophilic cysteine (Cys-$S_{nuc}$H) side chain is more exposed to the solvent and its interaction with the solvent can trigger the loss of proton to give the thiolate that was used to attack the central sulfur of the disulfide bond [64]. However, even though the way the proton from Cys-$S_{bur}$H is lost is yet to be fully understood [64,67], the presence of specific acidic groups like aspartate and glutamate near the cysteine are key for deprotonation [65,66]. This hypothesis may fail in cases in which the distance between the carboxylate of the Asp or Glu is more than 6 Å from the buried cysteine sulfur atom [64]. Based on this drawback, an alternative hypothesis of deprotonation using the anionic thiolate formed from the substrate after the first step has been suggested [67]. In a recent study of a similar enzyme protein disulfide isomerase (PDI) by Neves *et al.,* the hypothesis that the deprotonation occurs through a solvent-mediated proton transfer to the glutamate residue Glu-47 was suggested. [68].

The optimal orientation for the nucleophilic attack of the N-terminal active site cysteine involves an angle of 180 ° between the nucleophilic active site cysteine and the two sulfur atoms of the substrate (Fig. 3). This orientation results in a shortened projection of the disulfide bond of the substrate. Studies have revealed that changes of both the disulfide bond length and the dihedral angle are important prerequisites to increase efficiency of a nucleophilic attack [29,56,69].

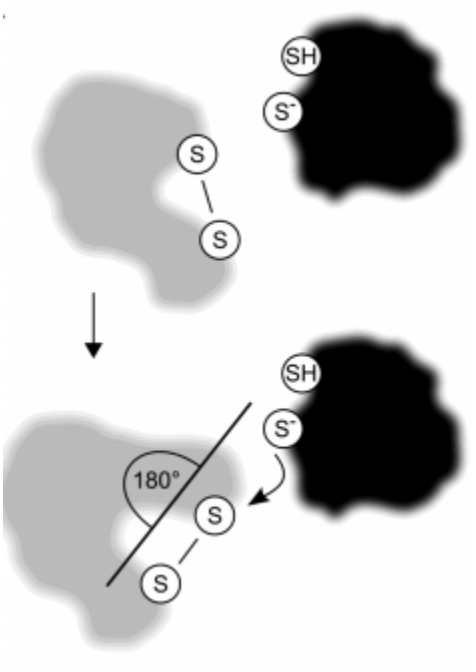



Fig. 3: Thiol-Disulfide exchange related structural changes [29].

The stabilization of the N-terminal cysteine anion thiolate has been proposed to arise from a variety of influences including the dipole moment of the active site helix [70-73], hydrogen bonding between active site thiols [74], hydrogen bonding between the N-terminal Cys sulfur atom and the amide proton of the C-terminal active site Cys [71,75-77], and electrostatic interactions between the thiolate anion and positively charged residues in the active site region [8,78]. Sun *et al.,* [53] performed theoretical calculations to address the factors that contribute to the stabilization of the N-terminal Cysteine using QUANTA97. The results revealed the existence of hydrogen bond interactions in nine or more of the conformers between the thiolate anion and the NH of Tyr25, the NH of Cys26, and the SH of Cys26 [53]. It was also reported that the side chain of the ammonium group of Lys20 was within 3.0 Å of the N-terminal Cysteine sulfur atom, implying that this residue may also contribute to the local positive electrostatic potential [53]. This outcome is similar to what was also observed by Jao *et al.,* in 2005, [79], but they went further to ascertain their conclusion by mutating Lys20 with leucine and glutamine. It was observed that the specificity of the Grx for glutathionyl was retained in both mutants, but their catalytic activity was lost due to higher pKa values for their N-terminal cysteine thiol moieties; thus, it was concluded that Lys20 contributed critically to the catalytic turnover [79].

The kinetic study of the human dithiol glutaredoxin as well as poplar GrxS12 shows that they have infinite $k_{cat}$ values, which are probably a common feature of enzymatically active Grxs [57,62,63,80] and glutathione-dependent hydroperoxides [81,82]. The rate constant of mammalian glutaredoxins and poplar GrxS12 ranges from $2.5x10^4$ to $2x10^6$ $M^{-1}S^{-1}$; this supports a rate limiting reduction half reaction [62,63,79,80].

# 1.4 Major Doubts

Researchers have done extensive work on glutaredoxins, as is well documented in the literature [25,33,38,43-48,51-55,57-63,76-81]. They have studied both the catalytic mechanism and the reaction kinetics of the protein and tried to analyze the factors that favor the active site motif anionic thiolate stabilization. In fact, some computational studies have analyzed the molecular basis for low pKa values at the N-terminal cysteine thiolate anion in hGrx and its role [79]. In addition, an explanation for the factors that are responsible for N-terminal cysteine thiolate anion stabilization was also put forward [79]. However, no computational work that specifically addresses the full catalytic mechanism of Grx was found in the literature; thus, this work is aimed at addressing this deficiency using human dithiol glutaredoxin. Hybrid quantum mechanics/molecular mechanics (QM/MM) methods were used in this study because they enable the study of large atomistic models, presenting results that involve the full enzyme [83]. In these methods, the QM region describes the part of the system where the reaction is taking place, that is, the active site, the substrate and other residues that are important for the reaction, while the MM part contains the rest of the enzyme [83].

# Chapter Two – Methods

## 2.0 Introduction

This chapter focuses on giving the background information on the computational methods as applied to biological systems (enzymatic reaction mechanism). The chapter contains four sections: (1) Quantum Mechanics (QM) Methods, (2) Molecular Mechanics (MM) Methods, (3) Hybrid Methods; particularly, QM/MM methods, and (4) Computational Enzymatic Catalysis.

The chapter starts with QM methods, which are known to give accurate results and capable of being used to study chemical reactions involving breaking and forming of chemical bonds. Different QM methods like *Ab initio*, Semi-empirical and Density Functional Theory methods are addressed.

In the second section (section 2.2), the concept of molecular mechanics (MM) methods, force fields and protein molecular force fields like AMBER, CHARMM and OPLS are discussed.

The third section dwells on hybrid methods with emphasis on QM/MM methods. This discussion covers the following main themes: the basis of QM/MM method, its types with more focus on the subtractive scheme (ONIOM methodology). Also, capping bonds at the QM/MM boundary, using both the link atom and the frozen atom approaches, and handling the non-bonded coupling terms using mechanical embedding, electrostatic embedding and polarized embedding schemes are discussed in the last part of this section.

Finally, the last section, section 2.4 provides information on computational enzymatic catalysis.

# 2.1 Quantum Methods

The study of enzyme catalysis via computational means requires method(s) that is(are) capable of describing the cleavage and the forming of chemical bonds. To explain this, it is important to have quantum mechanical (also referred to as electronic structure or quantum chemical) methods that can accurately predict the structures and the energetics of reacting groups in large molecular systems [84]. Quantum mechanics provides the conceptual framework for understanding chemistry and the theoretical foundation for a computational methodology that models chemical compounds through their electronic structure [85].

## 2.1.1 The Schrödinger Equation

The basis of quantum mechanics is the Schrödinger equation, which can give a complete description of the electronic structure of a molecule. If the equation could be fully solved, all the information relating to a molecule could be determined [86].

Quantum mechanics describes molecules in terms of interactions among nuclei and electrons, and molecular geometry in terms of minimum energy arrangements of nuclei [87]. All quantum mechanical methods ultimately trace back to the Schrödinger equation, which for the special case of hydrogen atom (a single particle in three dimensions) may be solved exactly.

$$[-\frac{1}{2}\nabla^2 - \frac{Z}{r}]\,\Psi(r) = E\Psi(r) \qquad \textbf{(6)}$$

The quantity in the square brackets represents the kinetic and potential energy of an electron at a distance of r from a nucleus of charge, Z (1 for hydrogen atom). The electronic energy in atomic units is the E and Ψ is the wave function that describes the motion of the electron as fully as possible and it is a function of the electron coordinates. The wave functions for the hydrogen atom are the conversant s, p, d atomic orbitals.

The Schrödinger equation cannot be solved exactly for molecular systems containing more than one electron; thus, approximations are needed for many-electron systems. The Schrödinger equation to a multinuclear, multielectron system can be generalized to:

$$\hat{H}\Psi = E\Psi \qquad \textbf{(7)}$$

The QM methods that are of interest for this thesis are, depending on the approximation made: the *ab initio*, semi-empirical and density functional theory (DFT) methods.

## 2.1.2 The Ab Initio Methods

*Ab initio* translated from Latin means “from the beginning”. It refers to the fact that no experimental data is used, and computations are based on quantum mechanics [88]. This is an approximate quantum mechanical calculation. The approximations made are usually mathematical approximations, such as using a simpler functional form for a function or finding an approximate solution to a differential equation [86].

The simplest and common *ab initio* calculation is the Hartree-Fock calculation, in which the mean field approximation is the main approximation. This means that the Coulombic electron-electron repulsion is not explicitly taken into account, however, its average effect is included in the calculation. This is a variational calculation, meaning that the approximate energies calculated are all equal to or greater than the exact energy. The approximation is the main flaw of HF as it ignores the tendency of electrons to avoid each other (electron correlation). This results in large errors that are associated with HF calculations and tend, with increasing basis size, to a limiting value called the Hartree-Fock limit [86].

In the Hartree-Fock method, the many-electron Schrödinger equation is broken into many simpler one-electron equations. Each one-electron equation is solved to yield a single electron wave function called an orbital, and an energy called an orbital energy. The orbital describes the behaviour of an electron in the net field of all the other electrons [86].

An additional issue that affects the accuracy of the computed results is the form chosen for the basis functions. The actual form of the single electronic molecular wave function (molecular orbital) is of course not known. The forms, used for the basis functions, can provide a better or worse approximation to the exact numerical single electron solution of the Hartree-Fock equation [86]. The basis functions used most often are combinations of either Slater type orbitals (exp (-ax)) or Gaussian type orbitals (exp ($-ax^2$)), abbreviated STO and GTO respectively. A molecular orbital is formed from a linear combinations of atomic orbitals, which are nothing more than linear combinations of basis functions with coefficients formed from the appropriate atomic HF calculations. Because of this approximation, most HF calculations give a computed energy greater than the Hartree- Fock limit [86].

Many correlated *ab initio* methods, such as those based on Møller-Plesset perturbation theory ($MP_n$, where n is the order of correction), configuration interaction (CI) or coupled cluster theory (CC), among others, use Hartree-Fock as a starting point. These methods offer a

substantial improvement in accuracy over Hartree-Fock calculations, but have a much higher computational cost, which makes their application for systems with tens of atoms difficult [84,86].

The Quantum Monte Carlo (QMC) is another way of avoiding making the Hartree-Fock mistake. There are several types of QMC: variational, diffusion and green function Monte Carlo calculations. These methods work with an explicitly correlated wave function and evaluate integrals numerically using a Monte Carlo integration. These calculations can be very time consuming, but they could yield extremely accurate results [86].

The merit of *ab-initio* methods is that they eventually converge to the exact solution, once all of the approximations are made sufficiently small in magnitude. However, this convergence is not monotonic [86]. Sometimes, a more approximate calculation gives a better result for a given property, than a more elaborate calculation. The demerit of *ab initio* methods is that they are expensive. These methods often take enormous amounts of computer CPU time, memory and disk space. The HF method scales as $N^4$, where N is the number of basis functions, so a calculation twice as big takes 16 times as long to complete. Correlated calculations often scale much worse than this. In practice, extremely accurate solutions are only obtainable when the molecule contains a few electrons.

In general, *ab initio* calculations give very good qualitative results and can give increasingly accurate quantitative results as the molecules in question become smaller [89].

### 2.1.3 Semi-Empirical Methods

Semi-empirical methods are the simplest electronic structure theory methods. They are the least computationally intensive of the quantum mechanical methods, but they are also typically the least accurate, unless specifically parameterized for a particular property [84]. The semi-empirical methods start out from the ab initio formalism and a drastic assumption to speed up the calculations is introduced by neglecting many of the less important terms in the *ab initio* equations. To compensate for the errors from these approximations, empirical parameters are incorporated into the formalism and calibrated against reliable experimental or theoretical reference data [90].

Semi-empirical methods can be applied to larger systems than DFT or correlated Hartree-Fock methods (typically hundreds of atoms) [84]. They can also be used in molecular dynamics simulations [91]. Semi-empirical approaches include: Modified Neglect of Diatomic

Differential Overlap (MNDO), Austin Model 1 (AM1), Parametric Model 3 (PM3), self-consistent charge density functional tight-binding (SCC-DFTB) etc.

AM1 and PM3 are based on exactly the same model as MNDO, and they differ from MNDO only in the implementation: the effective atom-pair potential in the core-core repulsion function is represented by a more flexible function with several additional adjustable parameters [92-94]. MNDO, AM1 and PM3 employ an sp basis without d orbitals in their original implementation. Based on this, they cannot be applied to most transition metal compounds [84]. The PM3 method uses the same formalism and equations as the AM1 method, but the differences include: PM3 uses two Gaussian functions for the core repulsion function, instead of the variable number used by AM1; the numerical values of the parameters are different. The other differences are based on the methodology used during the parameterization. AM1 takes some of the parameter values from spectroscopic measurements, while PM3 treats them as optimizable values [85].

The SCC-DFTB methods are based on density functional theory and have been shown to provide geometries and relative energies that are comparable to DFT and *ab initio* calculations [95]. SCC-DFTB methods are popular in the studies on biochemistry and material science [96,97] and their conceptual origin and derivation is quite different from those of the conventional semi-empirical methods, although the implementations and the actual computational procedures share many similarities [98]. SCC-DFTB methods can be considered as semi-empirical methods on par with the traditional ones because the methods employ severe integral approximation and extensive parameterization [98].

### 2.1.4 Density Functional Theory

Density functional theory can offer accuracy approaching that of the correlated *ab initio* methods but at substantial lower computational expense [84,99]. The DFT basis is that the ground state energy of a molecule can be calculated from the knowledge of the electron density distribution [84,100]. According to quantum mechanics, all the information needed about a given system is contained in the wave function. A wave function is a complicated function, for an N electron system, as it depends on 4N variables, i.e. one spin and three spatial coordinates for every electron (for fixed nuclear positions). The electron density is the square of the wave function and each spin density only depends on three spatial coordinates, $\rho(x,y,z)$. The electron density does not depend on the number of electrons; hence, it is independent of the size of the

system [101]. The use of electron density in DFT speeds up the calculation by simplifying the problem, while the wave function complexity increases exponentially with the number of electrons [102]. The basis for DFT is the proof by Hohenberg and Kohn that the ground state electronic energy is determined by electron density [102]. Numerous approximate functionals like Thomas-Fermi model, Hohenberg-kohn, Kohn-Sham, and so on, have been developed based on a mixture of trial and error and known limiting features of the exact functional, but there is no systematic way to improve them for now [84].

The Thomas-Fermi model of atoms is the first theory in which the wave function is not used to describe an atom. Instead, the electron density is used. It started by introducing a homogeneous electron gas called ‘’jellium’’ with a uniform positively charge background. This electronic distribution (called the uniform electron gas) has a constant non-zero density [101]. The Thomas-Fermi approximate functional has a number of mathematical flaws, such as wrong asymptotic behaviour of the electrostatic potential. Furthermore, this model predicts that the chemical molecules are always unstable. Thus, it is not applicable to chemical problems. It usually underestimates the kinetic energy of atoms and molecules by about 10% and its main flaw is the poor representation of the kinetic energy and the electron-electron interaction [101,103].

Hohenberg and Kohn stated a theorem about the importance of electron density. The Hohenberg-Kohn theorem states that all ground state properties of a system are determined by the density of the system. In this case, the total ground state energy of a many-electron system is a functional of the density and once the electron density is known, the total energy of the system is also known. Although this theorem ascertains the existence of a functional relating the electron density and the energy of a system, it does not tell the form of such functional. One of the main goals of DFT methods is to search for functionals that are capable of linking these two quantities [104].

Kohn-Sham approach as developed by Kohn W., and Sham L., in 1965 [105] introduced atomic orbitals, which is the basis of the current DFT application. This approach yields a way to solve the Hohenberg-Kohn theorem for a set of interacting electrons, starting from a virtual system of noninteracting electrons having an overall ground state density equal to the density of some real system of chemical interest where electrons do interact [102]. The idea in the Kohn-Sham formalism is to split the kinetic energy into two parts, one that can be calculated exactly and that considers electron as noninteracting particles, and a small correction term that

accounts for electron-electron interaction. The Kohn-Sham model is closely related to the Hartree-Fock method, sharing identical formulas for the kinetic, electron-nuclear and Coulomb electron-electron energies [101,103].

### 2.1.5 Exchange Correlation Functionals

The exchange-correlation functional accounts for the difference in the kinetic energy between the noninteracting system and the real system and also the difference between the classical and quantum mechanical electron-electron repulsion [101].

The hierarchy of density functional approximation can be described with Jacob's ladder (Fig. 4) [106]. The first introduced approximation was the local density approximation (LDA), and it is the simplest approach to represent the exchange-correlation function. LDA simplifies the electron-electron interactions by including the interaction between electrons and the charge density of the other electrons [107]. LDA is accurate despite its simplicity and it is mostly suitable for systems having slowly varying densities. Good results have been observed for several systems with relatively large density gradients [104]. On the other hand, LDA tends normally to underestimate atomic ground state energies and ionization energies, but it overestimates the binding energies [104]. The generalized gradient approximation (GGA) methods derive from the improvements of the local methods (LDA and LSDA). GGA functionals include the gradient, which makes the functional to contain their function of spin densities and their gradient [106,108]. GGA methods give better total energies [109], atomization energies [109-111], structural energy differences and energy barriers [112,113]. GGA methods give reliable results for hydrogen bridges, covalent, ionic, and metallic bonds, but they fail for van der Waals interactions [114,115]. Meta-GGA was developed based on GGA by including additional semi-local information beyond the first-order density gradient contained in the GGAs [104]. The hybrid density functional methods combine the conventional GGA method exchange-correlation with a percentage of Hartree-Fock or exact exchange [104]. These methods have become a popular choice in quantum chemistry and are now widely used [104].

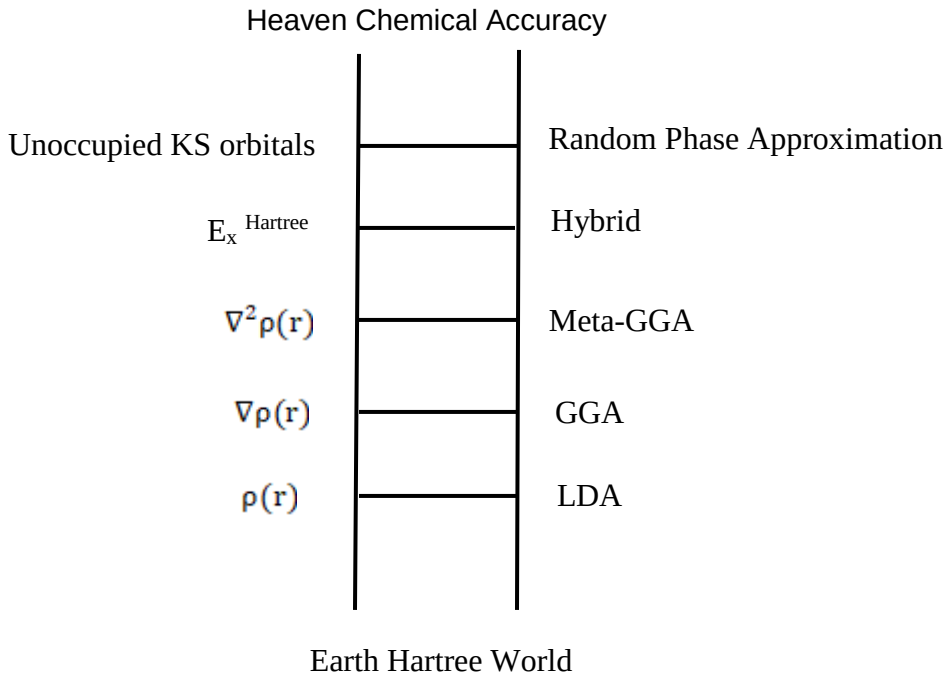



Fig. 4: Jacob's ladder for the DFT functionals [106].

### 2.1.6 Basis Sets

A basis set is a group of mathematical functions used to describe the shape of the orbitals in a molecule; each basis set has a different group of constants used in the wavefunction of the Schrödinger equation. The accuracy of a calculation is dependent on both the model and the type of basis set applied to it. There is a trade-off between accuracy and time. Larger basis sets will describe the orbitals more accurately but take longer to solve [89].

Basis sets for use in practical Hartree-Fock, density functional, Møller-Plesset and configuration interaction calculations make use of Gaussian-type functions. Gaussian functions are closely related to exponential functions, which are of the form of exact solutions to the one-electron hydrogen atom and comprise a polynomial in the Cartesian coordinates (x, y, z) followed by an exponential in $r^2$. Several series of Gaussian basis sets now have received widespread use and are thoroughly documented. Except for STO-3G and 3-21G, any of these basis sets can be supplemented with additional polarization functions and/or with diffuse functions. It should be noted that minimal (STO-3G) and split-valence (3-21G) basis sets, which lack polarization functions, are unsuitable for use with correlated models, in particular density functional, configuration interaction and Møller-Plesset models [116].

# 2.2 Molecular Mechanics

Molecular mechanics is the simplest and fastest way of evaluating molecular systems. It uses classical Newtonian mechanics to predict the energy of a molecule as a function of its conformation. In molecular mechanics, electrons are not explicitly included in the calculation, the energy of a system is calculated as a function of the nuclear position only. It can be used to supply the potential energy for molecular dynamics computations on large molecules, but they are not appropriate for bond cleavage and bond forming reactions. Molecular mechanics are faster than quantum mechanics and their computational requirements do not grow as fast with the size of a system.

Molecular mechanics is used in molecular dynamics to simulate the dynamic behaviour of molecules; in protein folding to predict the protein 3D structure from sequence; in protein-ligand docking to predict protein-ligand binding energy. Although molecular mechanics method is fast, it has many limitations: the method has limited precision that is significantly lower than QM calculations; it cannot be used to study chemical reactions that involves formation or cleavage of bonds; it requires parameterization.

## 2.2.1 Force Fields

Molecular mechanics uses force fields for the potential energy calculation between atoms. The force field energy is the sum of contributions due to covalent interactions and non-covalent interactions. A typical force field consists of bond stretching, angle bending, torsional rotation, van der Waals interaction and electrostatic interaction energy functions. Some force fields also include cross-terms. The final molecular mechanical energy, excluding cross-terms, is:

$$E_{MM} = E_{Stretching} + E_{Bending} + E_{Torsion} + E_{Electrostatic} + E_{VdW} \quad \textbf{(8)}$$

### 2.2.1.1 Bond Stretching

Bond stretching accounts for the deformation energies of bond lengths with respect to their equilibrium values. It originates from molecular spectroscopy, where bond displacement near the equilibrium position can be described using a harmonic potential.

$$E_{str} = \frac{K_b}{2}(b - b_o)^2 \quad \textbf{(9)}$$

where $K_b$ is the stretching force constant and is determined experimentally from Infra-red spectroscopy or estimated from quantum calculations, b and $b_o$ are bond length and equilibrium bond length respectively. The harmonic approximation can be used if atom bonding distances are near their equilibrium positions.

### 2.2.1.2 Angle Bending

The angle bending term describes deformation energies of the bond angles with respect to their equilibrium value. The energetic curve near the equilibrium angle can also be approximated by harmonic potential just like bond stretching. This approximation can also be used if bonding angles are near to their equilibrium positions.

$$E_{bend} = \frac{K_a}{2}(\theta - \theta_o)^2 \quad \textbf{(10)}$$

The bending force constants $K_a$ are also determined experimentally from Infra-red spectroscopy; alternatively, quantum calculations can be performed.

### 2.2.1.3 Torsional Terms

Torsions describe the rotation of the angle between planes through two sets of three atoms that have two atoms in common (dihedral angle) [117]. The origin of this energy is through space steric and electronic interactions between non-bonded atoms. Torsions can also take purely quantum-mechanical effects, like hyperconjugation, into account. The hyperconjugation plays a crucial role in many biological phenomena, for example, in influencing glycine conformations [118]. Torsions are useful in sampling conformational states of molecules in order to find local and global energy minima [117].

$$E_{torsional} = \sum \frac{V_n}{2}[cos(n\omega - \gamma) + 1] \quad \textbf{(11)}$$

### 2.2.1.4 Van der Waals Terms

Van der Waals term describes the van der Waals attractive and repulsive interatomic forces. This term is always approximated by the Lennard-Jones 12-6 potential, although other functions have also been proposed. The van der Waals term can be written as:

$$E_{vdW} = \sum_{i<j}^{N} \frac{A_{ij}}{r_{1j}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \quad \textbf{(12)}$$

where $A_{ij}$ and $B_{ij}$ are the Lennard-Jones parameters and $r_{ij}$ is the distance between atoms i and j.

### 2.2.1.5 Electrostatic Terms

Electrostatic terms describe the Coulomb interactions between atoms with partial charges. These terms are always difficult to calculate because they do not fall off rapidly with distance, and long range electrostatic interactions are often important features of the system under study, especially for proteins. The electrostatic term can be written as:

$$E_{ele} = \sum_{i<j}^{N} \frac{q_i q_j}{\epsilon r_{ij}} \qquad (13)$$

where $q_i$ and $q_j$ are atomic charges, $r_{ij}$ represents the distance between atoms i and j and $\epsilon$ is the dielectric constant.

## 2.2.2 Protein Molecular Mechanical Force Fields

Force field development requires the determination of parameters that are described in bond stretching, angle bending, torsional terms, van der Waals and electrostatic terms equations above. The applicability of molecular mechanics to a biochemical system is limited by the way these parameters were derived and for atom types [117]. The parameters are usually derived from experimental data [117]. The equilibrium angles, bond lengths and van der Waals radii can be determined from crystal structures [117]. Quantum mechanics calculations are often used in determining torsional potential and atomic charges because they are difficult to determine [117]. The main drawback of many of the existing force fields are inaccurate electrostatics and torsional potentials [117]. Numerous types of force fields are available: all-atom force fields, where parameters are considered for every atom; united atom force fields, where aliphatic hydrogen atoms are implicitly represented; and the coarse-grained force fields, where group of atoms are treated as super atoms.

In the description and study of proteins, Assisted Model Building with Energy Refinement (AMBER), Chemistry at HARvard Macromolecular Mechanics (CHARMM) and Optimized Potential for Liquid Simulations (OPLS) are the most commonly used molecular force fields. These force fields have same thing in common, that is, the potential energy function is a function of pairs of atoms.

### 2.2.2.1 CHARMM Force Fields

The first CHARMM force field was developed in the early 1980s by Martin Karplus and his group at Harvard [119]. The CHARMM force fields, which are a prominent set of force fields for studying biological systems, use classical and quantum mechanical energy functions for different types of molecular system. They include parameters for lipids, carbohydrates, nucleic acids and proteins. The united-atom, CHARMM19, the all-atom CHARMM22 and the dihedral potential corrected variant CHARMM22/CMAP are the CHARMM force fields designed for proteins [120,121]. CHARMM19 parameters aimed to provide a balanced interaction between water-water and solute-water energies. In CHARMM19 parameters, the hydrogen atoms bonded to nitrogen and oxygen were explicitly represented; hydrogen bonded to carbon or sulfur were treated as extended atoms [122]. A general version of the CHARMM force field (CGenFF) also exists and it allows the treatment of drug-like compounds while maintaining compatibility with other CHARMM force fields [123].

### 2.2.2.2 OPLS Force Fields

The Optimized Potential for Liquid Simulations was developed by William L. Jorgensen and his co-workers. These force fields are parameterized to simulate organic liquids and the properties of the liquid states of water [122]. There are different varieties of OPLS, that is, a united-atom and all-atoms versions, OPLS-UA and OPLS-AA respectively [123]. The OPLS-AA force field uses similar parameters as the Amber force field for bond angles and stretching [124]. The torsional parameters are obtained using data from *ab initio* molecular orbital calculations [124].

### 2.2.2.3 AMBER Force Fields

AMBER force fields were developed by Peter Kollman's group at the University of California, San Francisco for the calculations involving proteins and nucleic acids [125]. The AMBER force field contains values for the parameters of the corresponding equations like charges, force constants, equilibrium bond lengths and angles. The stretching and bending terms are calculated considering the harmonic approximation. The van der Waals interactions are approximated by the Lennard-Jones 12-6 formula, while the electrostatic interactions are calculated using the Coulomb formula. The torsional energies are determined by a Fourier series using the first six terms. The variations of AMBER force fields include: ff94, ff96, ff98,

ff99, ff99EP, ff02, ff02EP, ff03, ff10, ff12SB and ff14SB. The ff12SB reparameterizes backbone torsional angles, side chain torsions in selected amino acids, and incorporates improved backbone torsions in DNA and RNA [126]. ff14SB minimizes the dependency of protein side chain conformations on backbone conformations by including sidechain corrections and improves upon dihedrals in DNA and RNA. The General AMBER force field (GAFF) provides parameters for small organic molecules to facilitate simulations with drug-like molecules and small molecule ligands in conjunction with biomolecules [127]. Also, Rob Woods developed GLYCAM force fields (GLYCAM2000, GLYCAM04, GLYCAM04EP, GLYCAM06, GLYCAM06EP) for simulating carbohydrates. The AMBER energy function can be written as:

$$E = \sum_{bonds} \frac{K_b}{2}(b-b_o)^2 + \sum_{angles} \frac{K_a}{2}(\theta-\theta_o)^2 + \sum_{torsional} \frac{V_n}{2}[cos(nw-\gamma)+1] + \sum_{nonbond} [\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{\varepsilon r_{ij}}]$$

(14)

# 2.3 Hybrid Methods

The understanding of how enzymes facilitate chemical reactions is crucial, and atomistic simulations of enzymes can not only predict and establish their mechanisms but also provide valuable insight into the source of enzymatic rate enhancements [128]. To enable a sufficient balance in accuracy to describe chemical rearrangements and catalytic enhancement with the low computational cost needed to enable the inclusion of the full enzyme and/or sampling, a multilevel approach is employed [129-131]. Enzymes are too large to be described at any level of *ab initio* theory. At the same time, the available molecular mechanics force fields are not sufficiently flexible to model processes in which chemical bonds are cleaved and formed [132]. To overcome the limitations of a full quantum mechanical description on the one hand, and a full molecular mechanics treatment on the other hand, methods have been developed that treat a small part of the system at the level of quantum chemistry (QM), while retaining the computationally cheaper force field (MM) for the larger part [132]. This method is called a hybrid QM/MM. Also, hybrid QM/QM methods have been successfully used to study many enzymatic mechanisms [133-135]. In this method, a small part of the system is treated quantum mechanically using DFT or an *ab initio* method. This is often regarded as “high layer" and should include the parts of the system where bond cleavage and formation occurs. The larger non-reactive part, often named “low layer” comprises the remaining part of the system and it is always described with a simpler quantum mechanical method such as semi-empirical method,

AM1, PM3 or PM6. This hybrid QM/QM method is often used to study systems that contain between ~400 to ~700 atoms [136]. Although, both the QM/QM and the QM/MM methods have merits and demerits, a hybrid QM/MM methodology was used in this study owing to its popularity and its merit of being able to be used to model systems that contain the full enzyme.

## 2.3.1 QM/MM Methods

The hybrid QM/MM methodology was firstly introduced by Warshel and Levitt [130]. This methodology concurrently exploits the benefits of both the QM and the MM methods that were discussed in the last two sections. The region where the bond breaking and forming is taking place in the enzyme, along with selected catalytically relevant residues and water molecules, is treated with a QM method. The rest of the enzyme is treated with a MM force field as illustrated in fig. 5 below.

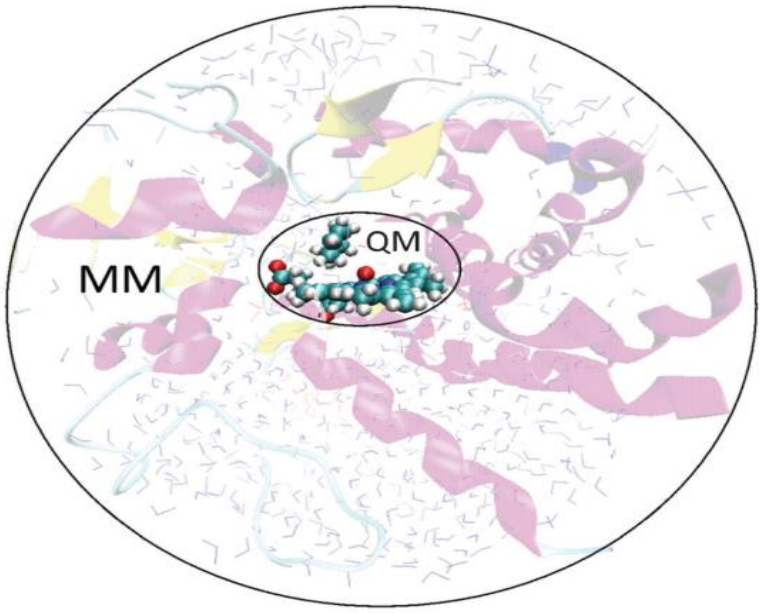



Fig. 5: Illustration of the QM/MM methodology. The system is divided into two layers: a small layer often regarded as the high layer, in which a chemical reaction occurs is treated with a QM method, while the rest of the enzyme and the surrounding solvent is modelled using MM [84].

The potential energy of the hybrid QM/MM comprises of three classes of interactions: interactions between atoms in the QM region, between atoms in the MM region and interactions between QM and MM atoms [84,132,136-139]. The QM and MM regions interactions are straightforward to describe using the QM and MM level respectively. However, the interactions between the two subsystems are more difficult to describe [132], and different approaches have

been proposed. These approaches can be roughly divided into two categories: subtractive and additive schemes [84,132,136-139].

## 2.3.2 TYPES OF QM/MM SCHEMES

The QM/MM methods are divided into additive methods and subtractive methods according to the way coupling between the QM and the MM part is performed.

### 2.3.2.1 Additive QM/MM Methods

In the additive scheme, the QM system is embedded within the larger MM system [132], and the total energy of the system is the sum of individual energies of the QM and MM regions, and the QM/MM coupling terms. The total energy of this type of scheme is expressed as:

$$E_{QM/MM,\ wholesystem} = E_{QM,highlayer} + E_{MM,lowlayer} + E_{QM/MM,coupling} - E_{QM/MM,correction} \quad \textbf{(15)}$$

The fourth term in the above equation is a correction term that is included to reduce the dependency of the total energy on the centers in the link atom region. The handling of this terms depends on the choice of the link atom coordinate, but in the majority of the additive schemes, this term is always neglected [136]. The major drawback with the use of this scheme is the difficulty in calculating the coupling term accurately, especially when link atoms are present and when electrostatics perturbations are included in the QM Hamiltonian [139].

### 2.3.2.2 Subtractive QM/MM Methods

In the subtractive scheme, the QM/MM energy of the system is obtained in three steps. First, the energy of the total system, consisting of both QM and MM regions, is evaluated at the MM level. The QM energy of the isolated QM subsystem is added in the second step. Third, the MM energy of the QM subsystem is computed and subtracted. The last step corrects for including the interactions within subsystem twice. The complete system energy expression is:

$$E_{QM/MM,\ wholesystem} = E_{MM,\ Real\ system} - E_{MM,\ Model\ system} + E_{QM,\ Model\ system} \quad \textbf{(16)}$$

The main advantage of the subtractive coupling scheme is that no communication is required between the quantum chemistry and molecular mechanics routines. This makes the implementation relatively simple. Unfortunately, this simplicity does come with a cost, as the coupling interaction between the high-layer and low-layer is handled at the low level of theory [140]. Also, there is no specific coupling term to take into account the interaction between the

QM and MM regions, as in additive schemes. On the other hand, a subtractive scheme requires a complete set of MM parameters for the inner subsystem, which may be difficult or cumbersome to obtain. Moreover, the coupling between the subsystems is handled entirely at the MM level. This is particularly problematic for the electrostatic interaction, which will then typically be represented by the Coulomb interaction between fixed atomic charges in the QM and MM regions [137].

Morokuma and co-workers have developed three methods that are based on the subtractive scheme: (1) the Integrated Molecular Orbital and Molecular Mechanics (IMOMM) method that integrates an orbital based method with an MM approach [141]; (2) the Integrated Molecular orbital and Molecular Orbital (IMOMO) method that combines two orbital based techniques [142]; and (3) the 'our own n-layered integrated molecular orbital and molecular mechanics' method (ONIOM) that allows the combination of an 'n' number successive of layers [143]. Recent improvements of the ONIOM approach [144,145] enable the inclusion of MM charges into the QM Hamiltonian (electrostatic embedding) and thus take it beyond a strictly subtractive scheme. The ONIOM method is a very popular choice nowadays [143], and it was employed in this study.

### 2.3.3 ONIOM Method

The development of the ONIOM method was born out of the improvements that were done in the IMOMM and IMOMO methods. These methods triggered many ONIOM methods improvements, as well as some alternatives for multilayer multilevel approaches. The IMOMM method can also be considered as the first generation of the ONIOM methods developed by Morokuma and co-workers. The IMOMM is a subtractive or extrapolative method [146] with the total energy as the one in equation 16. The subtractive method removes the "double-counted" MM contribution [141] as the subtraction of the MM energy of the real system from that of the model system is used to obtain the effect of the environment on the model system with an inexpensive MM method [146]. Thus, the IMOMM method can be considered as a size extrapolation from the accurate QM calculation of the small model system to the large real system by adding the approximate environmental effect evaluated by the inexpensive MM method [142].

Shortly after the IMOMM method development, Humbel *et al.,* extended the idea of the extrapolation step to a new hybrid method of combining two different QM (or MO) methods. This

method is the Integrated Molecular Orbital + Molecular Orbital (IMOMO) method [142]. The energy function for the IMOMO method is the same as the one expressed in equation 16.

The ONIOM method is regarded as a hierarchical method combining both the size of the system and the accuracy of the theory (or level) that approximates an accurate calculation for a large molecule [142]. In the original ONIOM scheme, all QM/MM interactions are treated classically by the MM force field, which is referred to as ONIOM-ME scheme [137]. An electrostatic embedding (ONIOM-EE) scheme was further developed to allow polarization of the QM wave function by the atomic charges of the MM part [145]. The ONIOM method does not need the additional coupling Hamiltonian when compared to the additive scheme. This makes the implementation of the ONIOM easier and avoids potential pitfalls of under- or overcounting the QM/MM interactions caused by a careless implementation of the coupling Hamiltonian. Furthermore, the ONIOM method can implicitly correct artefacts derived from the link atoms, if the MM potentials for the link atoms are similar to the QM potentials [137,145].

### 2.3.3.1 Capping Bonds at the QM/MM Boundary

The QM and the MM layers are generally connected by one or more covalent bonds when hybrid schemes are applied to enzymes. The presence of the covalent bonds brings some difficulties into the application of QM/MM methods to enzymes. In order to calculate the energies of each layer, it is necessary to disconnect the covalent bonds and generate two separate fragments [136]. In the literature, there are essentially two basic different classes of boundary schemes: (1) the link atom approach and (2) the frozen orbital approach. Both approaches have been used extensively and can result in similar levels of accuracy if they are used with care [139].

#### 2.3.3.1.1 The Link Atom Approach

In this approach, the boundary atoms between the QM and the MM layers are capped by the addition of an atom or a group of atoms along the chemical bond established between both regions, thus, mimicking the effect of the broken bond. Hydrogen is most often used, but other options such as hydrogen-like atoms, methyl groups, pseudohalogens can also be used to cap the QM subsystem [147]. Generally, the positions of the link atoms are chosen in such a way that they do not cut across any polar bond or unsaturated bond, avoiding any unrealistic effects. Each link atom introduces three additional degrees of freedom to the system, and despite this problem, link atoms are widely used. However, in practice the link atom is always placed at a

fixed position along the bond in every step of the simulation, so that these additional degrees of freedom are removed again [132].

### 2.3.3.1.2 The Frozen Orbital Approach

This approach of using a frozen orbital to saturate the dangling bond at the QM/MM boundary dates back to Warshel and Levitt [130]. The method assumes that the electronic structure of the bond is insensitive to changes in the QM region. The two most widely used approaches are the localized self-consistent field (LSCF) method which introduces orbitals at the QM atom and the generalized hybrid orbital (GHO) method which places additional orbitals on the MM atom.

The LSCF method was developed by Rivail and Assfeld [148]. In this method, the bonds between the QM and MM fragments are represented by strictly localized bond orbitals (SLBO), which are obtained by separate quantum mechanical calculations of small model compounds. These localized bond orbitals are assumed to be transferable for use in proteins and are kept constant throughout the SCF calculation. The LSCF method provides good results in energy minimization of reaction pathways in proteins and the assumption of transferability of bond orbitals appears to be a valid approach [149]. The LSCF method does not require the addition of link atoms into the system, although, the parameters for the localized bond orbitals have to be determined from model studies for each new system in the LSCF treatment [136].

The GHO developed by Gao and co-workers [150] is related to LSCF in that it constructs localized hybrid orbitals and freezes some of them [149]. The hybrid atomic orbitals are used as the basis functions on the boundary atoms of the MM fragment [150]. These hybrid orbitals are divided into auxiliary and active orbitals, the latter of which are optimized along with all other atomic orbitals of the QM layer fragment in the SCF calculations [136]. The chemical bond that connects the QM and the MM layers fragments is explicitly treated without introducing the link atoms [136].

The methods using local orbitals provide a better quantum mechanical description of the charge distribution around the QM/MM boundary when compared with the methods using link atoms [136]. The delocalized representation of the charges in these orbitals helps to prevent or reduce the over-polarization that is sometimes encountered in the link atom approach. However, the local orbitals methods are more difficult to implement than the link atom methods since they require several parameterization steps [147].

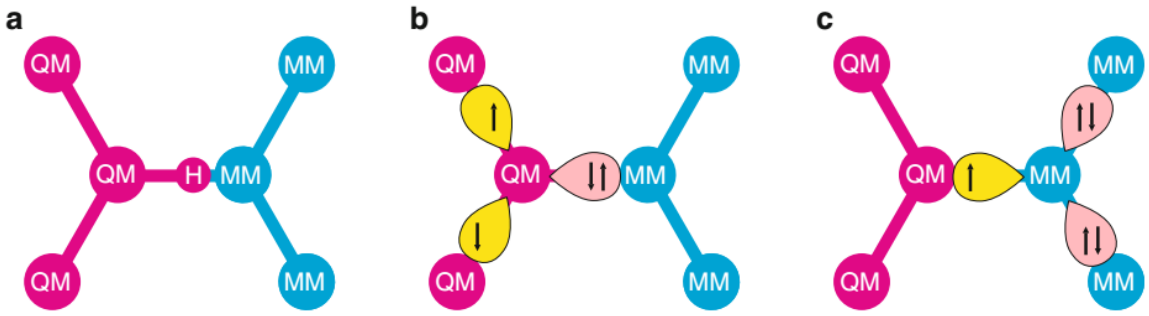



Fig. 6: Different methods to cap the QM region: link atoms (a) and frozen orbitals (b) and (c); b refers to LSCF orbitals and c to GHO orbitals [137].

## 2.3.3.2 Handling the non-bonded Coupling Terms

A very imperative issue in the hybrid QM/MM scheme is to deal with the non-bonded terms of the coupling between the high layer (QM) and the low layer (MM) regions. These interactions involve electrostatic as well as van der Waals forces that are necessary to be included in biological systems like enzymes [136]. Three different types of approaches can be considered for the electrostatic terms, and they include: mechanical embedding, electrostatic embedding, and polarized embedding.

### 2.3.3.2.1 Mechanical Embedding

The mechanical embedding approach is the simplest and the less computationally demanding method to treat the QM/MM electrostatic interactions in a classical manner. In this method, the QM calculation is performed in the gas phase, and it neglects any electrostatic effect of the environment in the QM region. The electrostatic interaction between the QM and MM regions is either omitted or in most cases, it is performed by the MM code, using a classical point charge model for the QM charged distribution [136]. In order to calculate the electrostatic interaction between the subsystems, one can either use a fixed set of charges for the QM region, for example, those given by the force field, or re-compute the partial charges on the QM atoms at every integration step of the simulation. The second strategy involves the use of a least-squares fitting procedure to derive atomic charges that optimally reproduce the electrostatic potential at the surface of the QM subsystem [151,152]. One of the drawbacks of

this approach is the need for accurate MM parameters for the QM region, and one common way of circumventing this involves the manual update of the MM charges for the QM region along the PES.

#### 2.3.3.2.2 Electrostatic Embedding

This approach includes the polarization of the QM region by the MM charge distribution in the QM electronic structure calculation. The partial charges used to describe the MM distribution are often taken to be those used in the force field relying on the use of electrostatic properties in the force field charge deviation. These are included in the QM Hamiltonian as an external potential of fixed point charges. This improvement makes the implementation of the electrostatic embedding scheme to be more difficult and leads to a growth in computational cost. However, it is important to take into consideration that this type of approach still neglects the less relevant polarization of the MM system by the QM region, an issue that can sometimes have a large effect in results [153,154].

#### 2.3.3.2.3 Polarized Embedding

The polarized embedding scheme further includes the polarization of the MM region in response to the QM charge distribution. In fact, both the QM and the MM regions should polarize each other until self-consistency is achieved in the charge distribution. These methods are difficult to implement and have the disadvantage of requiring well-established polarizable biomolecular MM force fields, which are yet to be available, but efforts are being conducted in their creation [155-158].

## 2.4 Computational Enzymatic Catalysis

Enzymes control many life processes and gaining a deeper understanding of enzyme catalysis is of great importance. A quantitative understanding of enzymes will not be possible without computer modelling approaches. Computational approaches have played a very crucial role in the understanding of the catalytic mechanisms of enzymes and have also provided better insights into the mechanism more than what can be observed from the experimental point of view, as it leaves many chemical questions unanswered.

Computational enzymatic catalysis is an interesting area of research that involves diverse studies targeted towards understanding the way enzymes work in certain circumstances

and to predict how they will behave in others. Predicting the catalytic mechanism of an enzyme is a complex and lengthy task. An essential first step in studying an enzyme-catalysed reaction is to establish the chemical mechanism. This means determining the identities of catalytic residues and other catalytic groups like cofactors and their roles and establishing the structures of any reaction intermediates and transition states.

Several computational methods, including Empirical Valence Bond (EVB), cluster modelling via first principles quantum mechanics (QM), Car-Parrinello Molecular Dynamics (CPMD) and hybrid quantum mechanics/molecular mechanics, among others, have been used extensively over the years to study the catalytic mechanisms of enzymes. The EVB method enables the determination of free energy profiles of chemical and enzymatic reactions with a remarkably low computational cost. EVB determines the activation free energy by incrementally changing the system adiabatically, from one diabatic state to another diabatic state [139]. EVB method makes it straightforward to use a non-geometrical reaction coordinate in modelling a reaction which may be significantly more accurate for some condensed phase reaction. The free energy surface can be calibrated by comparison with experimental data for reference reactions in solution in an EVB method. One great demerit of EVB method is that the initial mechanism of the reaction to be studied must be known before the method can be applied. Cluster modelling method uses only a model of the enzyme that contains a limited number of atoms. The selection of the amino acid residues to be included in the model is the main problem. All atoms that participate directly in the reaction must be included together with those that are involved in specific interactions at the active site. With the cluster model approach, it is possible to visually identify all structural changes in the model during a simulated reaction and to make certain that the changes are not artefacts of the chosen model [159]. The use of hybrid QM/MM method has been discussed in the latter section.

Computational enzymatic catalysis methods have the ability to determine and characterize the intermediates and the transition states from structural and energetic points of view, independently of their reduced lifetime and without interfering with the natural reaction flux [83]. Knowledge of how the transition states and the reaction intermediates are stabilized in an enzyme reaction can offer the potential to guide the design of high affinity drugs that mimic the transition states and thereby specifically inhibit their enzyme target [160].

# Chapter Three – Results and Discussion

## 3.0 Introduction

The chapter gives the account of the catalytic mechanism of glutaredoxin that was studied. The mechanism involves two basic steps: i) biomolecular nucleophilic attack ($S_N2$) of the N-terminal cysteine thiolate (Cys-24-thiolate) to the GSSG-disulfide to form the Grx-S-S-G mixed disulfide intermediate and ii) intramolecular oxidation via the deprotonated buried cysteine (Cys-27 thiol) to form the oxidized Grx (Grx-$S_2$) and the reduced glutathione (GSH). The catalytic scheme is shown in Fig. 7. The chapter contains the computational methods, results and discussion as well as the conclusions drawn from each mechanistic step. The computational methods are the model preparation, building the QM/MM model and the determination of the catalytic mechanism sections of this chapter. The results and discussion section present the results and the analysis of the results obtained for each step with their respective conclusions. The results revealed that the first thiol-disulfide exchange (first step) has higher Gibbs activation energy than the second one (7.8 vs 5.9 kcal.mol$^{-1}$), implying that the second step is much faster than first step. The studies predict the first thiol-disulfide exchange (first step) to be the rate limiting step as the rate of reaction can only be measured from the slowest step.

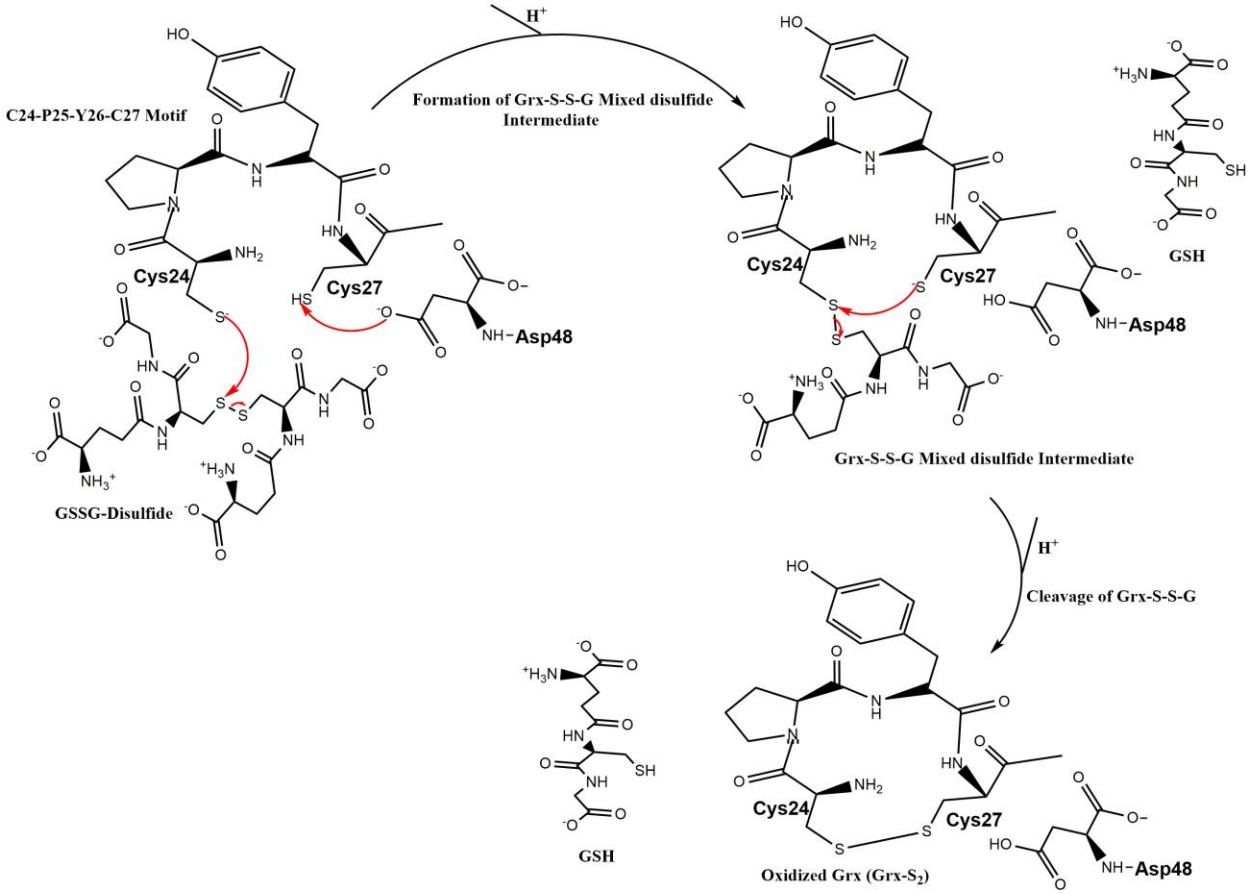



Fig. 7: Catalytic Mechanism Scheme for the reduction of GSSG-disulfide with Grx [26,28,29,33,57,63,64].

## 3.1 Study of the Step One: Biomolecular Nucleophilic Attack ($S_N2$) of the Cys-24-Thiolate to the Central Sulfur atom of the GSSG-Disulfide to form the Grx-S-S-G Mixed Disulfide Intermediate.

The Grx-S-S-G mixed disulfide intermediate formation in this step involves the nucleophilic attack of the Cys-24-thiolate to the central sulfur atom ($S_{ctr}$) of the GSSG. It is a redox reaction to reduce the GSSG to the mixed disulfide intermediate in which the enzyme (Grx) is covalently bound with the substrate (GSSG-disulfide) through a new disulfide bridge (Fig. 8).

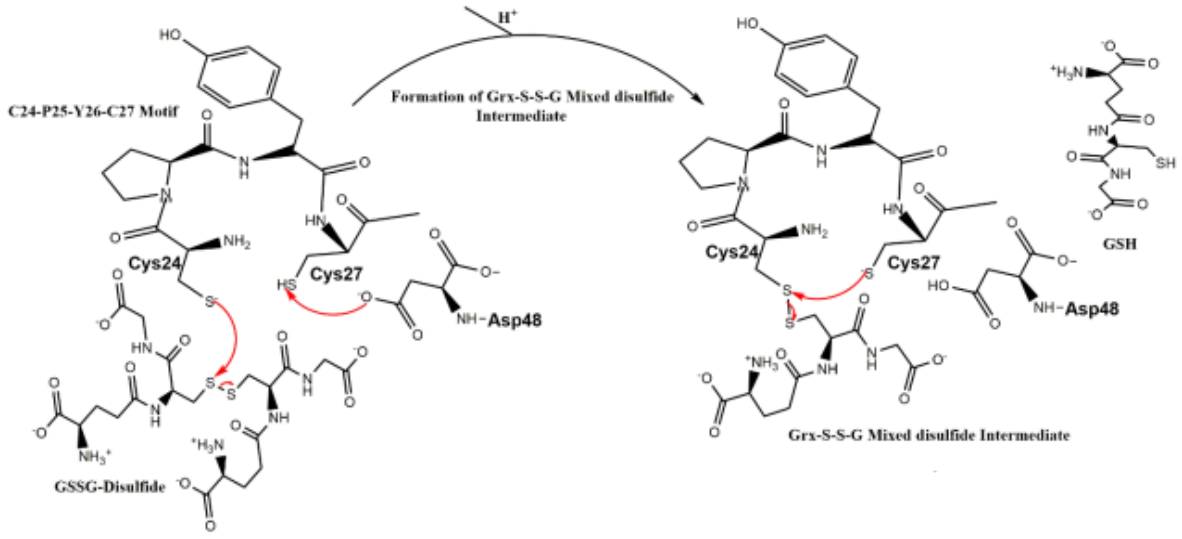



Fig. 8: First Step Reaction showing the formation of Grx-S-S-G Mixed disulfide Intermediate.

## 3.1.1 Model Preparation

### 3.1.1.1 Modelling Glutathione Disulfide (GSSG) into the Human Glutaredoxin

The X-ray crystallographic structure of human glutaredoxin (hGrx) (PDB ID code: 4RQR, 1.08 Å resolution) [161] was used for this study. The GSSG substrate was modelled in the active site of the hGrx from the X-ray structure of another glutaredoxin of bacterial origin (PDB code: 4TR0, 1.95 Å resolution) [162]. The two X-ray structures were first aligned, and the alignment gave a RMSD value of 0.73 Å (Fig. 9). The software X-leap [163] was used to protonate the complex, assuming the physiological protonation states for all residues and H++ web-server software was used to validate the protonation state [164]. Cys24 (N-terminal cysteine) was modelled as a thiolate because it is solvent exposed with low pKa value. One $Na^+$ counterion was added with X-leap to compensate for the negative charge of the system. The system was surrounded by a cubic box solvated with Transferable Intermolecular Potential 3-Point (TIP3P) water molecules within a distance of at least 12 Å from the surface of the enzyme.

A four-stage minimization of the geometry using MM was performed in order to eliminate clashes and bad contacts, prior to MD. In the first stage, only the water molecules were minimized, with the rest of the system being fixed with 50 $kcal.mol^{-1}Å^{-2}$ harmonic constraints. In

the second stage, all the hydrogen atoms of the system were optimized, with the heavy atoms being restrained with a 50 kcal.mol$^{-1}$Å$^{-2}$ harmonic potential. In the third minimization stage, only the CA and N backbone atoms of the protein and the modelled hGrx were fixed. Finally, in the last stage, all atoms were geometry-optimized with no restraint.

Molecular Dynamics (MD) simulations were run with the Amber software [163] so that the solvent can be equilibrated to the minimized X-ray conformation of the protein. All the atoms in the modelled complex except the ones from Asp48, Cys27 and the atoms from the tail part of the GSSG (substrate) were restrained with 20 kcal.mol$^{-1}$Å$^{-2}$ harmonic force constants. The system was heated during 100 ps in the canonical (NVT) ensemble by linearly increasing the temperature from 0 K to 310 K. Afterward, a MD simulation within an Isobaric-isothermal (NPT) ensemble at a fixed temperature and pressure of 310 K and 1 bar respectively was run without any restraint for 2 ns (Langevin thermostat, integration step of 2 fs, non-bonded interaction cut-off distance of 10 Å). The SHAKE algorithm was used to constraint the bond length of those bonds involving hydrogen atoms. The trajectories were saved every 1000 steps for analysis. One structure from the MD trajectory was selected as the representative catalytic conformation of the solvated hGrx:GSSG complex. The structure was selected based on: i) closeness to linearity of the bond angle between the Cys-24 sulfur atom (N-terminal cysteine) and the two sulfur atoms of the GSSG-disulfide (substrate); ii) shortness of bond distance between hydrogen atom of the Cys27 (C-terminal cysteine) SH group and the carboxylate oxygen atom of Asp48 that was hypothesized to cause the deprotonation of the hydrogen.

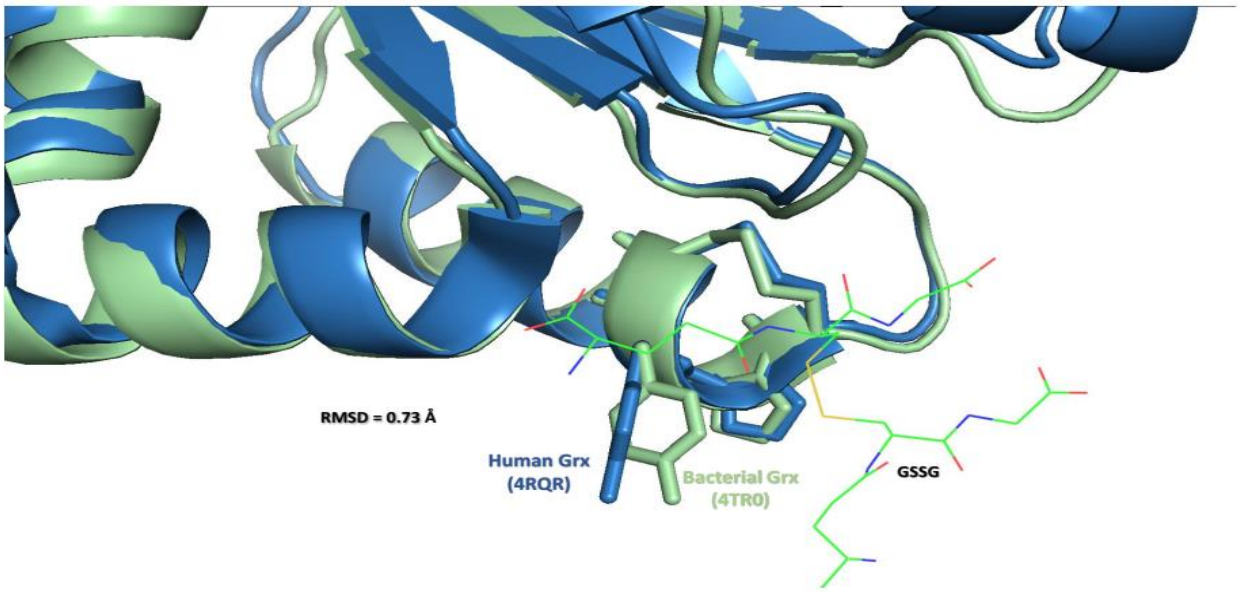



Fig. 9: Alignment result of the active site of the human glutaredoxin (light blue) and the bacteria glutaredoxin (pale green).

### 3.1.1.2 Building the QM/MM Model

The QM/MM method with our own N-layered integrated molecular orbital/molecular mechanics (ONIOM) methodology as implemented in Gaussian09 package was employed in this study. The system was partitioned into two layers: a high layer that was treated at the DFT level, and a low layer, treated at the classical MM level. The QM/MM model contains the hGrx, the GSSG, and all the water molecules within 6 Å of the hGrx:GSSG complex, summing up to 5,132 atoms (Fig. 10). The high layer described at the DFT level was comprised of 137 atoms (with overall charge of -2 and singlet multiplicity) and included the Cys24-Pro25-Tyr26-Cys27 motif; Asp48 sidechain; a significant part of GSSG (with charge of 0); some selected water molecules (5 water molecules) that are within 5 Å from the GSSG sulfur atoms and the carboxylate oxygen atom of Asp48 (Fig. 11). The low layer contained the remaining part of the model (4,995 atoms). Hydrogen atoms were used as the link atoms to complete the valences of the bonds spanning between the low and the high layers of the ONIOM QM/MM scheme. Another model that contained 125 atoms in the DFT layer was also explored to know the effect of not including the whole peptide chain of both Cys24 (N-terminal cysteine) and Cys27 (C-terminal cysteine) in the DFT layer. Also, a similar model used only to study the first step of the reaction was also prepared. This model contained 99 atoms (overall charge of -1 and singlet multiplicity) in the DFT layer and included the Cys24-Pro25-Tyr26-Cys27 motif; an important part of GSSG (with charge of 0); one water molecule from the X-ray crystallographic structure that is within 4 Å from the GSSG sulfur atoms. This model was purposely created to assess if the inclusion of the water from the X-ray crystallography structure alone in the DFT layer was enough to stabilize the leaving glutathione.

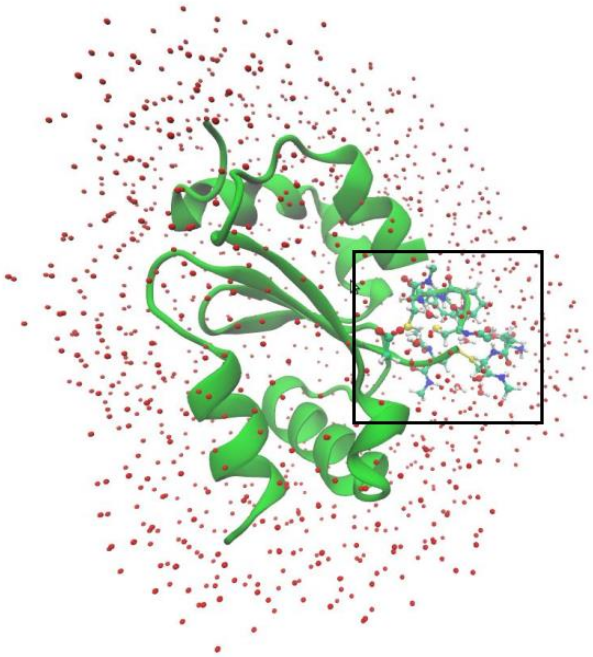


Fig. 10: The optimized hGrx:GSSG complex with the water molecules represented as red dots.

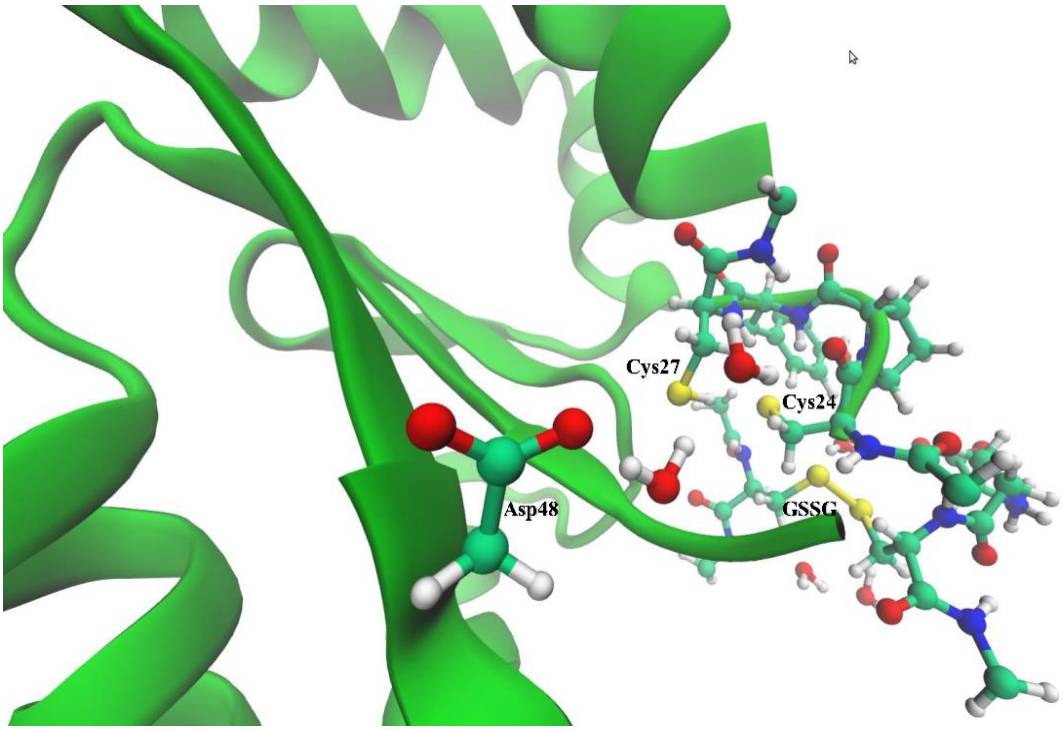



Fig. 11: The DFT layer that contained the $C^{24}P^{25}Y^{26}C^{27}$ motif, Asp48 sidechain, important part of GSSG and five water molecules.

### 3.1.1.3 Determination of the Catalytic Mechanism

The calculations were performed with the ONIOM methodology and the mechanical embedding scheme was employed for the description of the Coulomb and van der Waals interactions between the QM and MM regions using the Gaussian09 package. The mechanical embedding scheme was resorted to because the reaction could not be studied after several attempts when the electrostatic embedding scheme, which includes the polarization of the QM region by the MM charge distribution in the QM calculation was used. The mechanical embedding scheme has been extensively used with success in the study of reaction mechanisms of various enzymes [165,166,167]. Throughout the calculations, all the water molecules in the MM layer were fixed to their original positions, while all the enzyme residues and substrate atoms (in both QM and MM layers) were kept free. The low layer was described with the Amber ff99SB force field [163] and the high layer was accounted for with DFT and the mPW1N functional. This functional was chosen because previous studies have shown that it is one of the best functionals to adequately describe the thermochemistry and kinetics of thiol-disulfide exchange [167]. The 6-31G(d) basis set was used at this point, because it has been shown to be adequate for the geometry optimization [139].

After the full geometry optimization of the three models that were explored, linear transit scans along the reaction coordinate ($S_{nuc}$ – $S_{ctr}$) for this step were carried out with 0.05 Å increments to locate the transition state. The scans results are presented in figures below (Fig. 12-15). The transition state was characterized and its full optimization as well as frequency calculation were carried out using the same level of theory and basis set used for both the initial geometry optimization and the scans.

The 0 K electronic energy was determined with single-point energy calculations at ONIOM(mPW1N/6-311+G(2d,2p):Amber level of theory for the reactant, the transition state and the intermediate. The zero-point energy (ZPE), thermal and entropic energy corrections were calculated using mPW1N/6-31G(d):Amber at 298.15 K and 1.0 bar during the frequency calculation of each minima and TS, in order to build the thermochemical and kinetic profile of the reaction. Gaussian provides a thermochemistry section in the frequency calculation output, which has data about ZPE, and thermodynamics energies, enthalpies and Gibbs free energy [168].

## 3.1.2 Results

The transition state could not be characterized in the two models with 99 atoms and 125 atoms in the DFT layer. In the smallest model (with 99 atoms in the DFT layer), the energy increased throughout the scan (Fig. 12) and this implies that the presence of the X-ray crystallographic water alone in the DFT layer is not enough for the reaction to occur. The result of the model with 125 atoms in the DFT layer, that is aimed at examining the effect of not including the whole Cys24 and Cys27 peptide chains in the DFT layer, also followed similar trend as the one with 99 atoms in the DFT layer (Fig. 13).

The scan of the model that contained the whole Cys24 and Cys27 peptide chains and more water molecules from the MD simulation (with 137 atoms in the DFT layer) gave a maximum energy structure along the distinguished reaction coordinate that corresponds to the rearrangement of the DFT layer water molecules arising from a greater increase in the low layer energy (Fig. 14). However, the energy scan also showed a stable intermediate. The reverse scan that was performed along the $S_{ctr}$ - $S_{lg}$ coordinate starting from the relaxed intermediate structure resulted in a transition state that is characteristic of thiol-disulfide exchange (Fig. 15). From the obtained results, it can be concluded that: i) the inclusion of the whole Cys24 and Cys27 backbones in the QM layer is necessary to study this mechanism; ii) the presence of only a water molecule from the X-ray crystallographic structure in the QM layer is not enough to study the mechanism, and there is need to include more water molecules from the MD simulation in the DFT layer. Based on these findings, the model with 137 atoms in the DFT layer was used for the study because it solved all the drawbacks identified in the previous models.

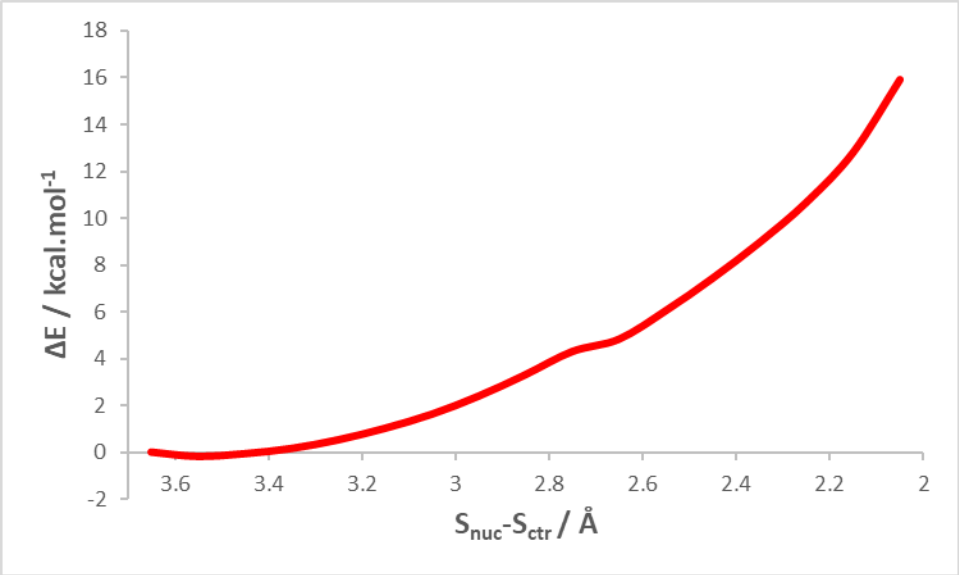



Fig. 12: The scan plot for the model with 99 atoms in the DFT layer.

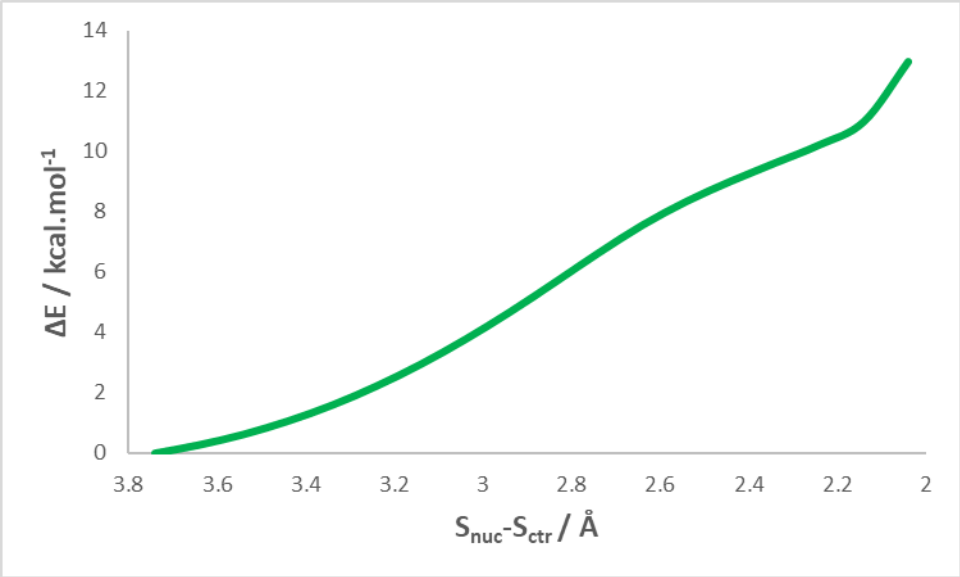



Fig. 13: The scan plot for the model with 125 atoms in the DFT layer.

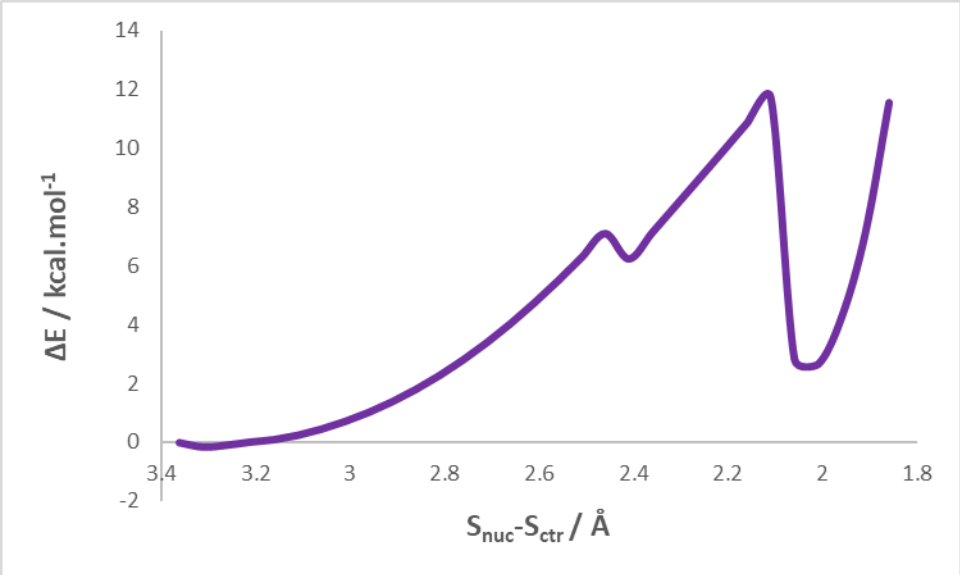



Fig. 14: The scan plot for the model with 137 atoms in the DFT layer.

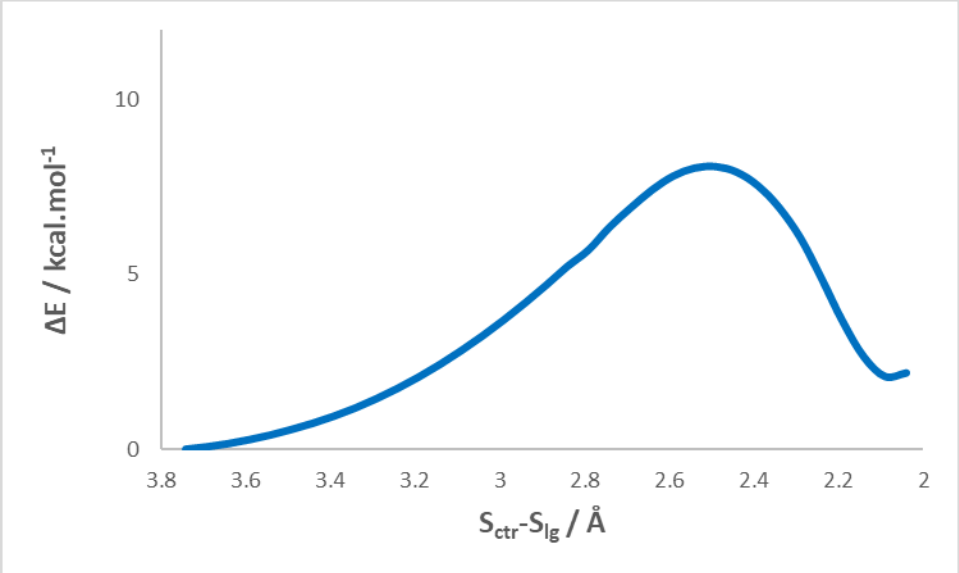



Fig. 15: The plot for the reverse scan of the intermediate from Fig. 14.

## 3.1.3 Analysis and Discussion of the Results

In the optimized QM/MM model, the GSSG sulfur atoms form an angle of 163.5° with the Cys24-sulfur atom, this is close to the linearity required for the thiol-disulfide exchange reaction, as the cleavage of disulfide bond is thought to occur without the participation of 3d orbitals [169]. The backbone NHs of Tyr26 and Cys27, as well as the SH group of Cys27 are positioned to favour the formation of hydrogen bonds with Cys24-thiolate (distances being 2.45 Å for

Tyr26(NH), 2.48 Å for Cys27(NH) and 2.40 Å for Cys27(SH)). These interactions are responsible for the lower pKa value of Cys24 as they stabilize the negative charge of thiolate. The Asp48 carboxylate, which is hypothesized to deprotonate Cys27-SH makes a hydrogen bond with the water molecule used as a bridge to mediate the proton transfer at a distance of 1.74 Å at both the reactant and the transition state of the first step. Furthermore, all angles that encompass the GSSG sulfur atoms ($S_{ctr}$ and $S_{lg}$) or Cys24 sulfur atom ($S_{nuc}$) are close to 90° ($C_{ctr}$-$S_{ctr}$…$S_{cys24}$ is 91.43°, $S_{ctr}$…Scys24-C is 80.66°, $C_{lg}$-$S_{lg}$-$S_{ctr}$ is 105.69°, and $C_{ctr}$-$S_{ctr}$-$S_{lg}$ is 104.30°). Also, all the equilibrium dihedral angles around the S-S axis are close to 90° (99.11° for $C_{ctr}$-$S_{ctr}$-$S_{cys24}$-C and -80.30° for $C_{lg}$-$S_{lg}$-$S_{ctr}$-$C_{ctr}$), all these values are in good agreement with the values found and proposed in the literature [68,169,170].

The substrate (GSSG) is well solvent-exposed as seen in Fig. 10. When only the water molecule from the X-ray crystallographic structure was used (model with 99 atoms in the DFT layer), no transition state was found (Fig. 12). This result is similar to what was observed in a study of a homologous enzyme (protein disulfide isomerase, PDI) by Neves *et al.,* [68]. However, upon inclusion of more water molecules from the MD simulation with the full peptide chain of Cys24 and Cys27 (model with 137 atoms in the DFT layer), a transition state with a low activation barrier of ~7.8 kcal.mol$^{-1}$ was observed. These results buttress the point that water molecules play a very crucial role in the nucleophilic attack of Cys24 to the central sulfur atom of the GSSG disulfide, as it was reported in previous studies [56,64,68]. The low barrier implies that the reaction is very fast, which is in agreement with what was reported in the literature [57,171]. The comparison of this barrier with the one of homologous enzyme, PDI, (18.7 kcal.mol$^{-1}$) studied by Neves *et al.,* [68] reveals that the nucleophilic attack of N-terminal cysteine to the central sulfur atom of GSSG-disulfide in Grx is faster than the one in PDI. This result is as expected, because the cyclic sidechain of proline (pyrrolidine) in the $C^{24}P^{25}Y^{26}C^{27}$ motif of Grx activates the N-terminal cysteine, $C^{24}$, making the nucleophilic attack to occur faster, unlike PDI (with active site motif $C^{53}G^{54}H^{55}C^{56}$), where the glycine has hydrogen as its sidechain.

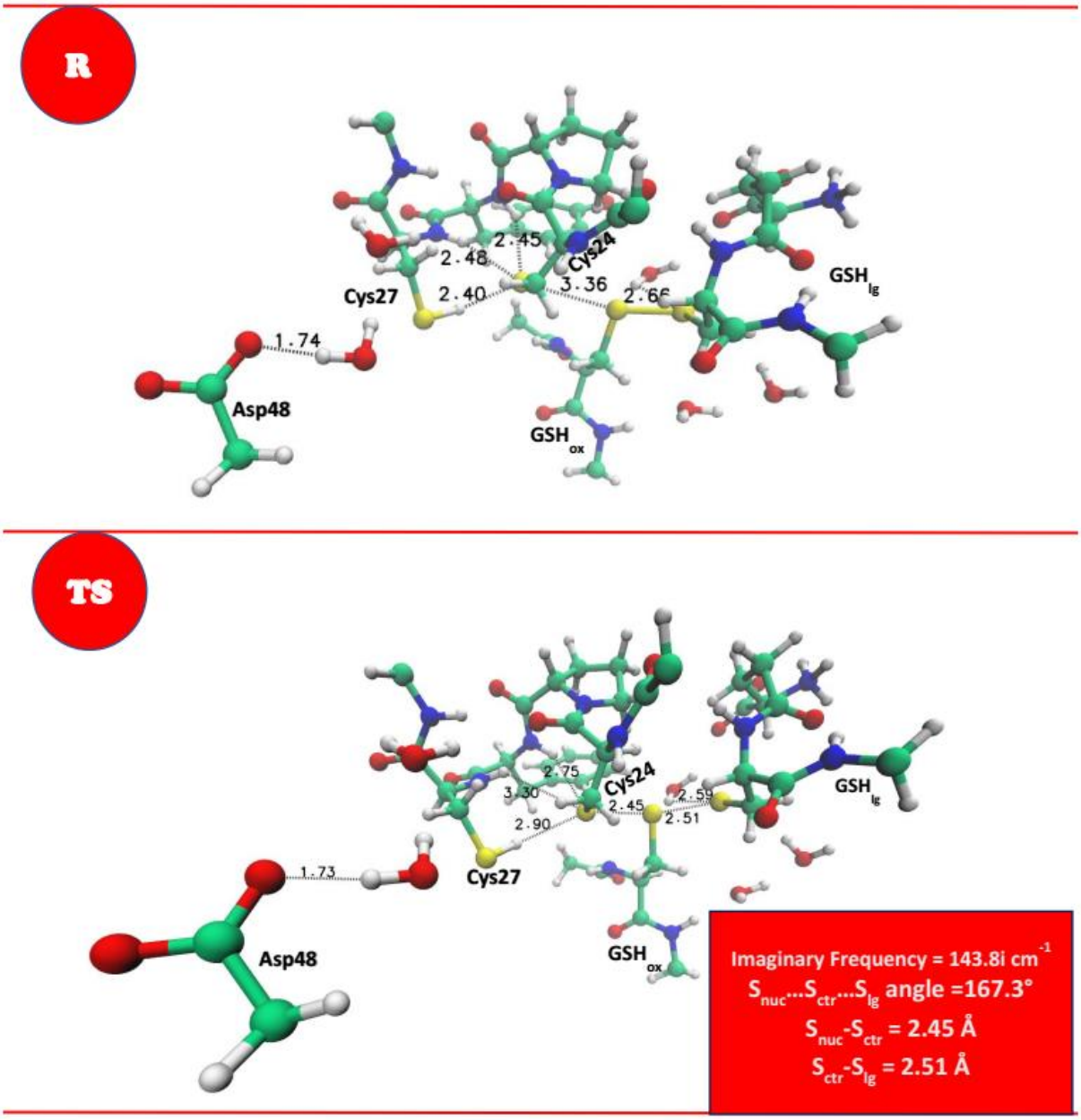
Fig. 16 shows the relevant chemical structure for the first thiol-disulfide exchange (first step of the process). In the reactant, the central sulfur atom of GSSG and the Cys24 sulfur atom were at a distance of 3.36 Å, and the Cys24 thiolate was stabilized via hydrogen bond formation with backbone NHs of Tyr26 and Cys27 and also by the SH group of Cys27.

The nucleophilic attack of the Cys-24-thiolate to the central sulfur atom of the GSSG gave the expected mixed disulfide intermediate in which the Grx and GSSG are covalently bound via a new disulfide bridge. The full optimization of the TS structure and vibrational frequencies analysis showed one imaginary frequency of 143.8i $cm^{-1}$, that coincided with the reaction coordinate. This vibration displayed a linear antisymmetric stretching of the three sulfur atoms with little contributions from the backbone NHs (Tyr26 and Cys27), SH (Cys27) and the water molecule that stabilized the leaving glutathione. The trisulfide anion ($S_{nuc}…S_{ctr}…S_{lg}$) exhibited a bond angle of 167.31° and bond distances of 2.45 Å for $S_{nuc}$-$S_{ctr}$ and 2.51 Å for $S_{ctr}$-$S_{lg}$. These values are typical of a thiol disulfide exchange reaction transition state where a linear arrangement of the three sulfur atoms is required [56,59,64,68]. One of the water molecules that is in the vicinity of the leaving glutathione makes a hydrogen bond (2.59 Å) with the anionic sulfur atom of the leaving group and thus stabilizes the leaving glutathione anion. Also, the hydrogen bond networks between the Cys24-thiolate and the backbone NHs of Tyr26 and Cys27, as well as the SH group of Cys27 showed larger distances at the TS than at the reactant state, 2.75 Å, 3.30 Å and 2.90 Å, respectively. These increments are because the Cys24-thiolate sulfur is becoming less nucleophilic in preparation for the formation of the new disulfide bond between the Grx and the GSSG.

At the mixed disulfide intermediate state, the hydrogen bonds between the Cys24-thiolate and Cys27 SH group, as well as with the backbone NHs of Tyr26 and Cys27, increased when compared to what was observed at both reactant and transition state. Furthermore, a charge transfer from Cys24-thiolate sulfur ($S_{nuc}$) (q values for $S_{nuc}$ are -1.11 a.u. and -0.42 a.u. at reactant and intermediate states respectively) to the leaving group glutathione sulfur ($S_{lg}$) (q values for $S_{lg}$ are 0.74 a.u. and -0.79 a.u. at reactant and intermediate states respectively) was observed due to the breakage of the disulfide bond of the GSSG substrate.

The Gibbs activation energy, including zero-point energy contributions, for this first step is 7.8 $kcal.mol^{-1}$ and it is much lower than the experimental result of 17.2 $kcal.mol^{-1}$ calculated from the pseudo-first order rate constant of 1.66 $s^{-1}$. This value represents only an upper limit to which the theoretical values must obey [171]. From the thermodynamics parameter table (Table 1), the Gibb energy and the enthalpy have similar behaviour. The entropy contributions were very small as seen in Table 1 below, thus this reaction was enthalpy driven which is similar to the result obtained by Neves *et al.,* for a homologous enzyme (Protein Disulfide Isomerase). The insignificant contribution of entropy to this reaction is normal due to the bond breaking and bond formation process associated with this step. Also, the reaction free energy of the first step

is -0.6 kcal.mol$^{-1}$, implying that the step is exergonic since the intermediate state is slightly more stable (has lower energy) than the reactant state.



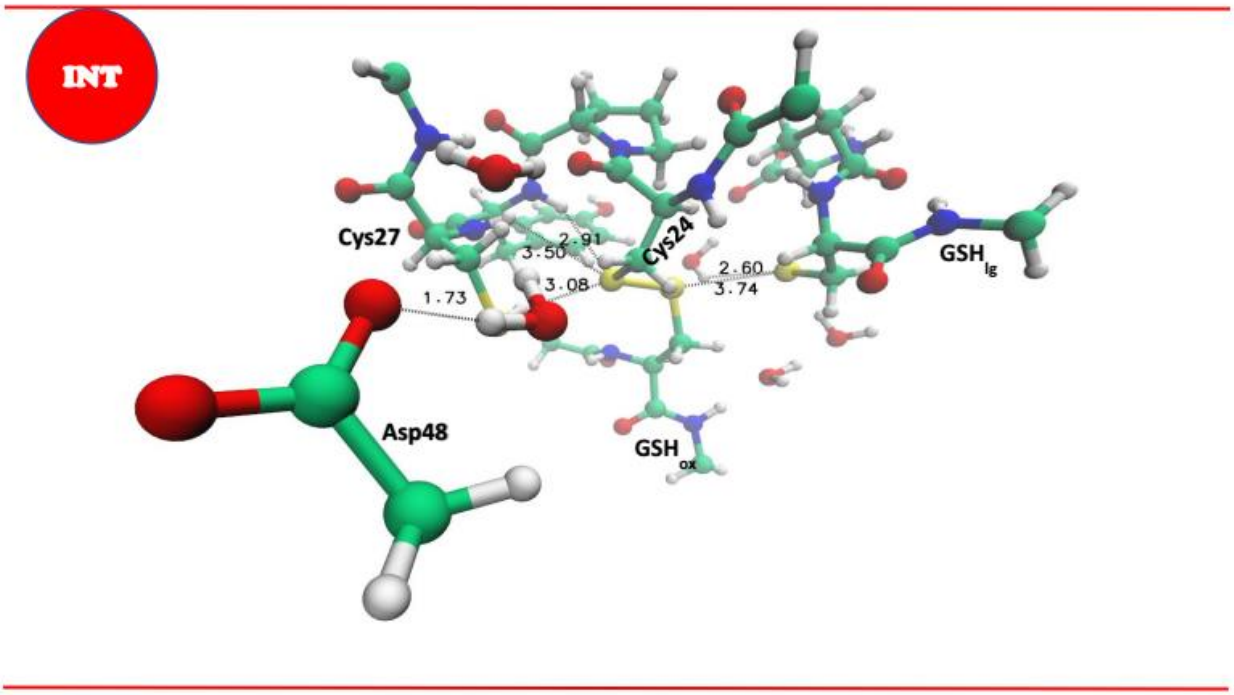



Fig. 16: The reaction states for the nucleophilic attack of the Cys24-thiolate to the GSSG-disulfide to form the mixed disulfide intermediate (step1).

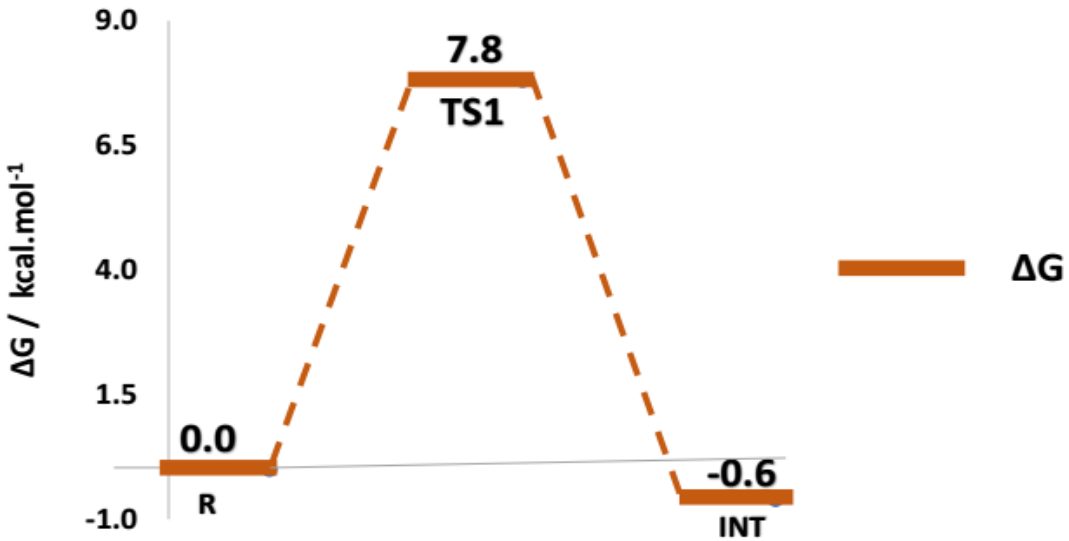



Fig. 17: Thermochemical Profile for the First Step.

Table 1: Thermodynamics Parameters Values for the First Step

| | ΔG / kcal.mol$^{-1}$ | ΔH / kcal.mol$^{-1}$ | -TΔS / kcal.mol$^{-1}$ |
|---|---|---|---|
| Reactant | 0.0 | 0.0 | 0.0 |
| Transition State | 7.8 | 7.8 | 0.0 |
| Intermediate | -0.6 | 0.1 | -0.7 |

### 3.1.4 Conclusions from the First Reaction Step

The first step involves the biomolecular nucleophilic ($S_N2$) attack of the N-terminal cysteine thiolate (Cys-24 thiolate) to the central sulfur atom of the substrate (GSSG-disulfide) to form the so-called mixed disulfide intermediate in which the enzyme (Grx) is covalently bound with the substrate (GSSG-disulfide) through a new disulfide bridge. The results of the calculations reveal that water molecules play a vital role in this nucleophilic attack and that the consideration of just a water molecule from the X-ray crystallographic structure alone is not enough to reproduce the reaction. Also, the results show that the inclusion of the whole peptide chain of both Cys24 and Cys27 and more water molecules from the MD simulation in the DFT layer are important in the study of the mechanism.

The transition state gave a trisulfide anion ($S_{nuc}…S_{ctr}…S_{lg}$) with bond angle of 167.31° and bond distances of 2.45 Å for $S_{nuc}$-$S_{ctr}$ and 2.51 Å for $S_{ctr}$-$S_{lg}$, which are characteristics of a thiol-disulfide exchange reaction transition state. The TS gave a low activation barrier of approximately 7.8 kcal/mol, which is both thermodynamically and kinetically favourable. This result implies that the nucleophilic attack of the Cys-24 thiolate to the central sulfur atom of GSSG is very fast, which is in agreement with what was reported in the literature.

## 3.2. Study of the Step two: Intramolecular Nucleophilic Attack of the Deprotonated Cys27 on Cys24 to form an oxidized Grx and the reduced Substrate.

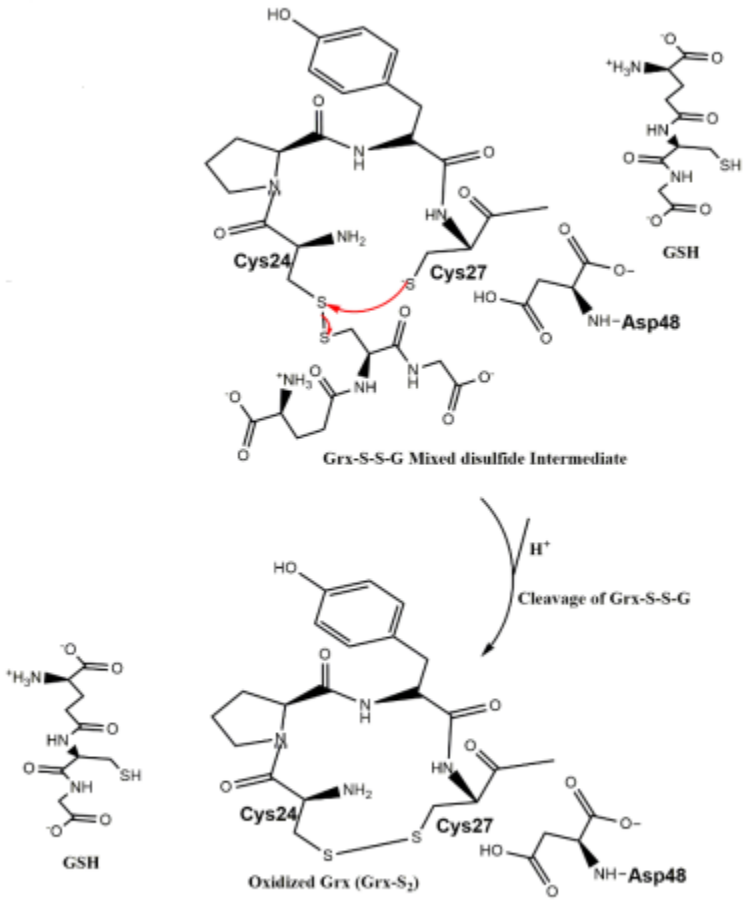



Fig. 18: Second Step Reaction showing the formation of the oxidized Grx (Grx-$S_2$) and the second reduced GSH.

This is the second thiol-disulfide exchange reaction. Here, the Cys24 of the mixed disulfide intermediate is nucleophilically attacked by the deprotonated C-terminal cysteine (Cys27) to form the oxidized Grx and the reduced dithiol substrate (Fig. 18). Prior to this attack, the buried cysteine (Cys27) in the mixed disulfide intermediate formed in the first step was deprotonated by Asp48. This proton transfer was mediated by a water molecule as proposed by

Menchise et al., [65] because Asp48 and Cys27 are more than 6 Å apart (~9 Å). The transfer of the buried cysteine proton to the Asp48 sidechain carboxylate is expected to be a very fast process with very low barrier [65,67,68], but in this study, a very high barrier of approximately 25 kcal.mol$^{-1}$ was obtained for this proton transfer (see Fig. 19). This barrier is too high for a proton transfer that is expected to be fast. The product formed after the deprotonation is more stable (has lower energy) than the starting reactant, which is an indication that the process is thermodynamically favourable, but not favourable kinetically (Fig. 19). The second thiol-disulfide exchange reaction was studied based on the assumption that the buried cysteine (Cys27) was deprotonated by the Asp48, since the deprotonation stage is not a rate determining step in the reaction. However, the detailed mechanism through which the buried cysteine is deprotonated to perform the intramolecular oxidation is still yet to be understood [64,67].

### 3.2.1 Determination of Second Step Mechanism

The deprotonated mixed disulfide intermediate was employed as starting point for this second step (intramolecular oxidation). In this intramolecular oxidation, three different models were tested. The first model has the negatively charged $GSH_{lg}$ group (leaving glutathione) removed from the QM layer (Fig. 20), and the second model has the $GSH_{lg}$ group protonated and retained in the QM layer (Fig. 21). In the third studied model, the negatively charged $GSH_{lg}$ group was retained in the QM layer without being protonated. For this mechanistic step, linear transit scans along the reaction coordinate (Cys27-$S_{bur}$…$S_{nuc}$-Cys24) were performed. The reaction coordinate scan was performed using 0.05 Å increments and these increments were decreased to 0.01 Å when near to the transition state. The obtained transition state was fully optimized, and the frequency calculation of the fully optimized TS was done using the same level of theory and basis set as used in the first step. The presence of one imaginary frequency from the vibrational frequencies calculation confirmed the nature of the TS structure.

The energies for the optimized geometries of step 2 starting reactant, the TS and the product were obtained by single point calculation using the mPW1N functional and a larger basis set (6-311+G(2d,2p)). The frequency and thermochemical calculations for each of the reaction states were then carried out at 298.15 K and 1.0 bar to obtain the ZPE, entropic and thermal contributions to the free energy. These frequency calculations were done using the mPW1N functional with the 6-31G(d) basis set, as in initial geometry optimizations and in the linear transit scans.

## 3.2.2 Results

The scans along the reaction coordinate for the first two models, where the negatively charged leaving glutathione ($GSH_{lg}$) group was removed from the QM layer and where it was protonated and retained in the QM layer, revealed no transition states as the energy increased throughout the scan from the reactants to the corresponding products states (Fig. 20 and Fig. 21). The third model with the negatively charged $GSH_{lg}$ group retained in the QM layer gave a maximum in the energy profile and thus a candidate for the transition state of this reaction (see Fig. 22). This result indicates that the intramolecular oxidation through the deprotonated Cys27-thiol would rarely occur when the product of the first step is still bound to the active site. It further indicates that the repulsion that would arise from the thiolates of the negatively charged glutathione groups of both the products of the first and second steps have no significant effect on the formation of the oxidized Grx. A molecular dynamics simulation of the mixed disulfide intermediate with $GSH_{lg}$ negatively charged was run for 10 ns. The result of the simulation (Fig. 23) shows that the negatively charged $GSH_{lg}$ was drifting away from the active site during the course of the simulation, and this supports the point that the intramolecular oxidation would occur when the product of the first step has already left the active site.

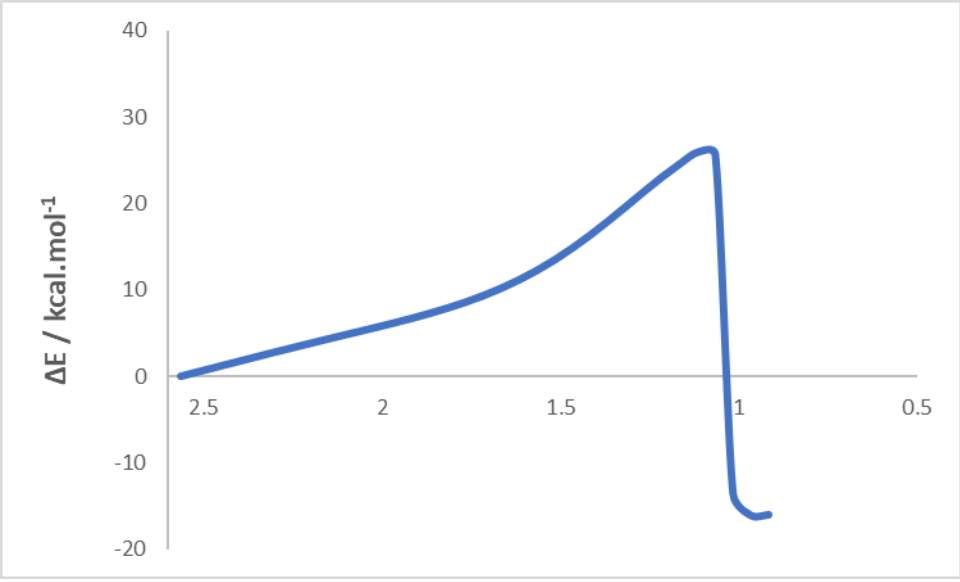



Fig. 19: The scan plot for the deprotonation of Cys27-thiol by Asp48 via a bridging water molecule.

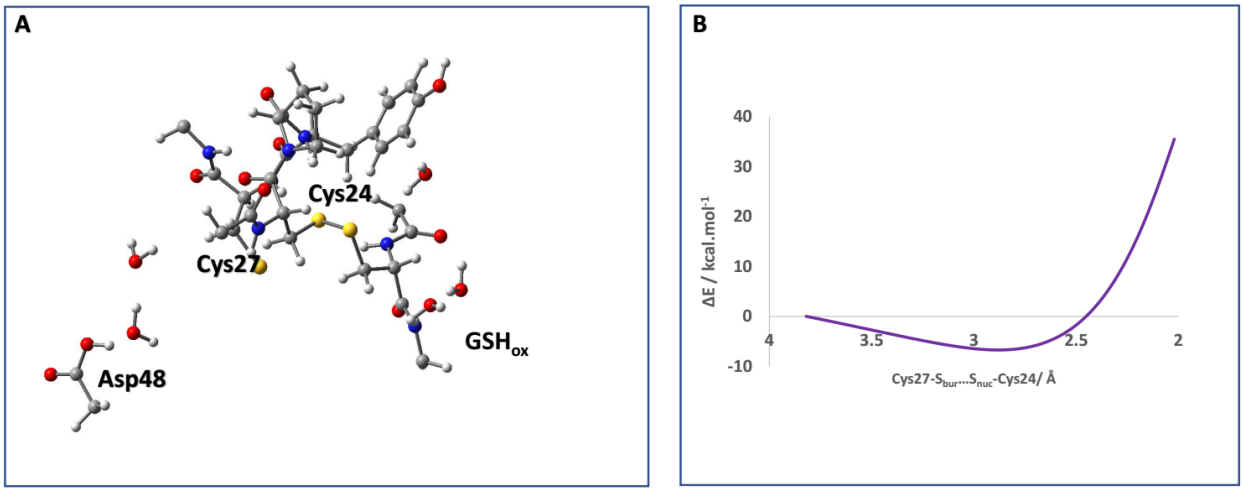



Fig. 20: Structure (A) and Scan plot (B) for the model with $GSH_{lg}$ removed from the QM layer.

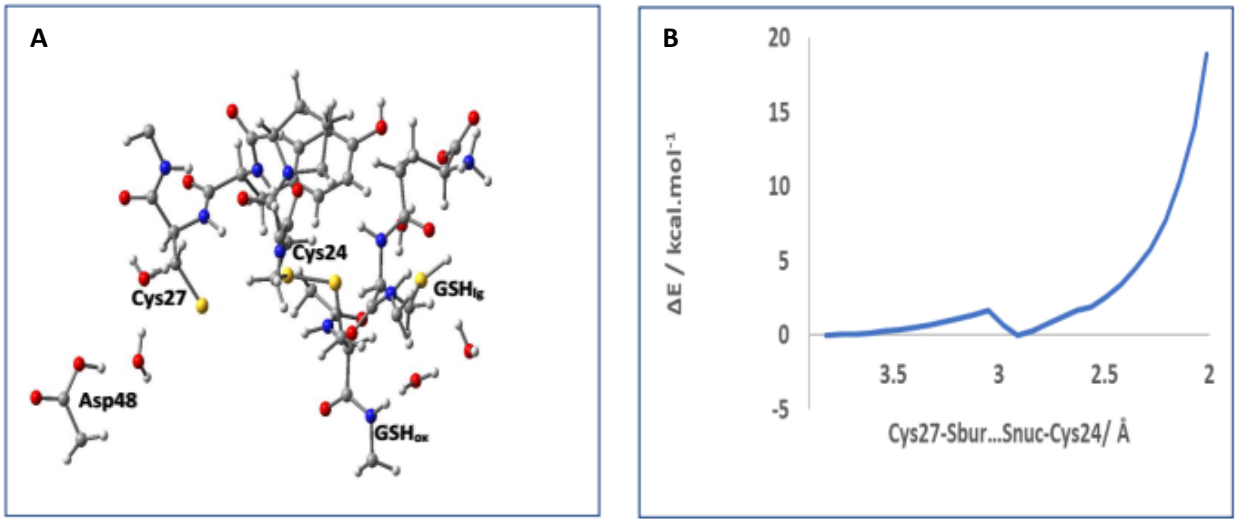



Fig. 21: Structure (A) and Scan plot (B) for the model with $GSH_{lg}$ Protonated and retained in the QM layer.

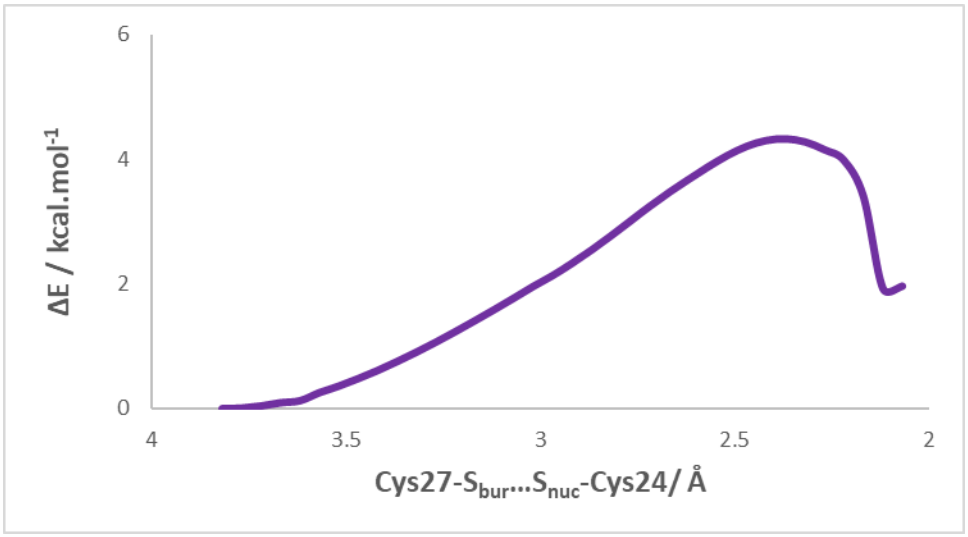



Fig. 22: The scan plot for the intramolecular oxidation of the Grx-SSG Intermediate.

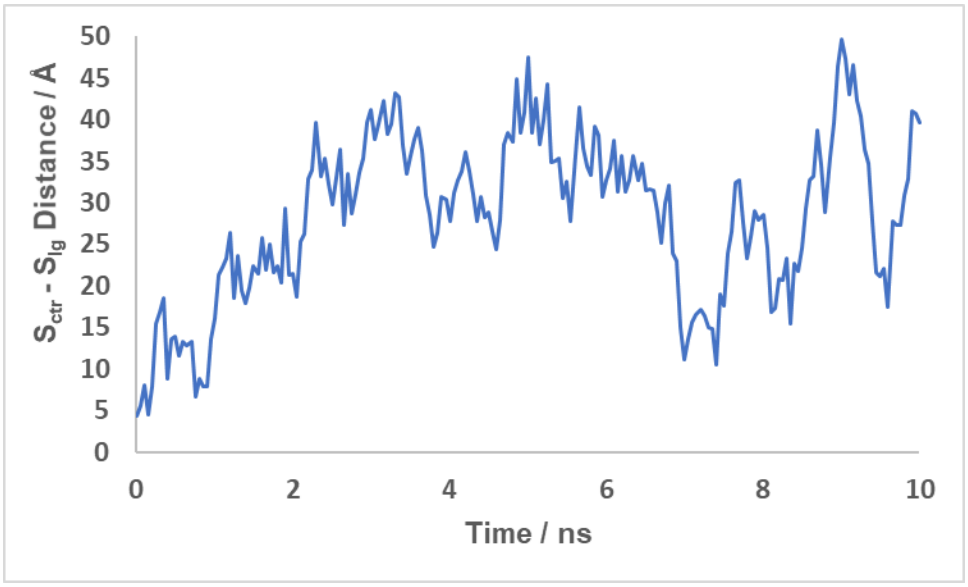



Fig. 23: The MD simulation plot of the Grx-SSG Intermediate showing how the distance between the thiolates of $GSH_{ox}$ and $GSH_{lg}$ changed with time during the simulation.

### 3.2.3 Analysis and Discussion of the Results

The reaction takes place in a similar way as the first step, but with smaller rearrangement of water molecules around the glutathione moieties, which also showed a smaller number of hydrogen bond interactions when compared with the first step. In the starting reactant (Cys27-thiol deprotonated mixed disulfide intermediate), the Cys27 and Tyr26 amide backbones are still orientated to favour the hydrogen bond formation with the sulfur atom of Cys24 (2.84 Å for Tyr26 and 3.19 Å for Cys27). The Cys-27 thiolate is placed in a linear arrangement with the mixed disulfide sulfur atoms, forming an angle of approximately 163°. The O-H of one of the water molecules in the vicinity of the negatively charged $GSH_{lg}$ group was stretched (0.96-0.99 Å) to favour the stabilization of the leaving glutathione moiety.

At the transition state, a barrier of 5.9 $kcal.mol^{-1}$ was obtained. The hydrogen bond interaction between the Cys24 sulfur and the backbone NH of Cys27 was shortened (3.19-2.67 Å), while the one of Tyr26 re-orientated toward the central sulfur atom ($S_{ctr}$) of $GSH_{ox}$. The analysis of the vibrational frequencies of the optimized TS geometry revealed only one imaginary frequency of 81.1i $cm^{-1}$, corresponding to the antisymmetric stretching of the three sulfur atoms ($S_{bur}$, $S_{nuc}$ and $S_{ctr}$) with minor contributions from Cys27 and Tyr26 backbone NHs groups. The three sulfur atoms ($S_{bur}$…$S_{nuc}$…$S_{ctr}$) exhibited a bond angle of 169.64° which is close to the linear arrangement (180°) required for thiol-disulfide exchange reaction. In addition, bond distances of 2.38 Å and 2.58 Å were obtained for $S_{bur}$-$S_{nuc}$ and $S_{nuc}$-$S_{ctr}$ respectively. These values are also typical of the symmetric transition states that are always observed in thiol-disulfide exchange reaction [56,59,67,68].

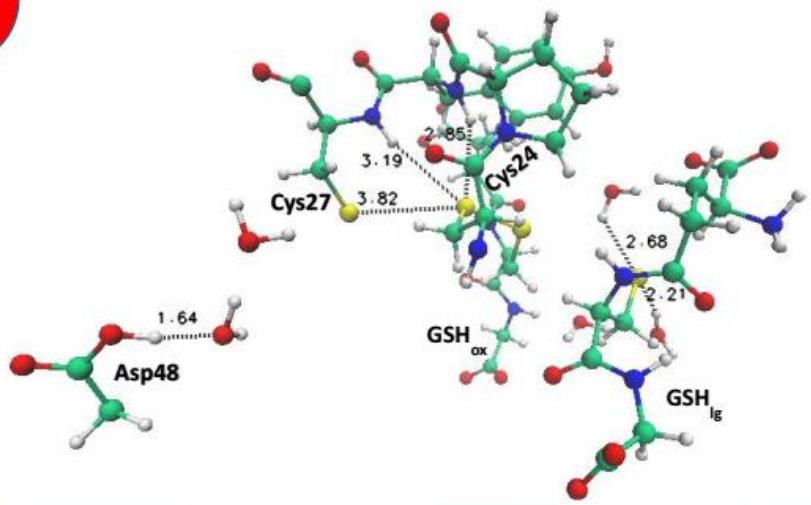
At the product state, the expected oxidized Grx is formed via a new disulfide bond between the N and C termini cysteine sulfur atoms in addition to the two molecules of reduced glutathione. The new S-S bond length is 2.09 Å. These reduced GSHs are negatively charged and are stabilized by the nearby water molecules. The two GSHs are widely separated (~5 Å) due to the repulsive interaction between the two as they are both negatively charged (see product state in Fig. 24). One of the water molecules acts as a bridge between the two GSHs, forming two hydrogen bonds (2.39 Å for $GSH_{red}$ and 2.71 Å for $GSH_{lg}$). Another water molecule in the vicinity of $GSH_{lg}$ also forms a hydrogen bond with the thiolate (2.18 Å).

Furthermore, there was charge transfer from the deprotonated Cys27-thiolate sulfur ($S_{bur}$) (q values for $S_{bur}$ are -0.95 a.u. and 0.05 a.u. at the reactant and the product states,

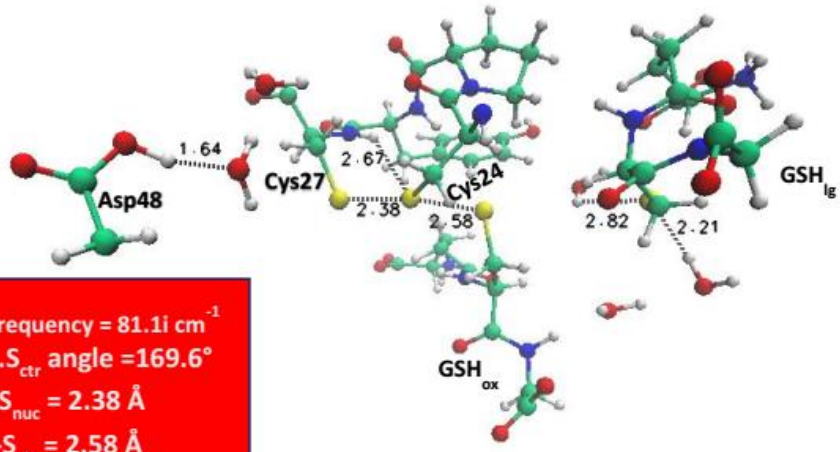
respectively) to the reduced glutathione sulfur ($S_{ctr}$) (q values for $S_{ctr}$ are 0.43 a.u. and -0.50 a.u. at the reactant and the product states, respectively). This charge transfer is due to the cleavage of the mixed disulfide S-S bond and the formation of new S-S bond between the active site Cys27 and Cys24 residues.

The Gibbs activation energy of the second thiol-disulfide exchange reaction (second step) is 5.9 kcal.mol$^{-1}$ which is lower than the first thiol-disulfide exchange (first step). This implies that the second thiol-disulfide exchange is faster than the first thiol-disulfide exchange, suggesting that the first step is the rate determining step since it is the slowest of the two steps. There is no available experimental result in the literature to compare the result of this step with. The activation as well as the reaction Gibbs energies are mediated by the relative stabilization of the leaving anionic thiolates by the environment [68]. The reaction free energy is 1.6 kcal.mol$^{-1}$, suggesting that the product is slightly endergonic. However, this value (1.6 kcal.mol$^{-1}$) is within the error of the method pointing to a reaction that is almost thermoneutral, and if the release of the product would be considered, the free energy would be lower because of the entropic increase. Figures 24 and 25 show the reaction states and the thermochemical profile for the second step respectively.

**INT**




**TS**




Imaginary Frequency = 81.1i $cm^{-1}$

$S_{bur}...S_{nuc}...S_{ctr}$ angle =169.6°

$S_{bur}$-$S_{nuc}$ = 2.38 Å

$S_{nuc}$-$S_{ctr}$ = 2.58 Å

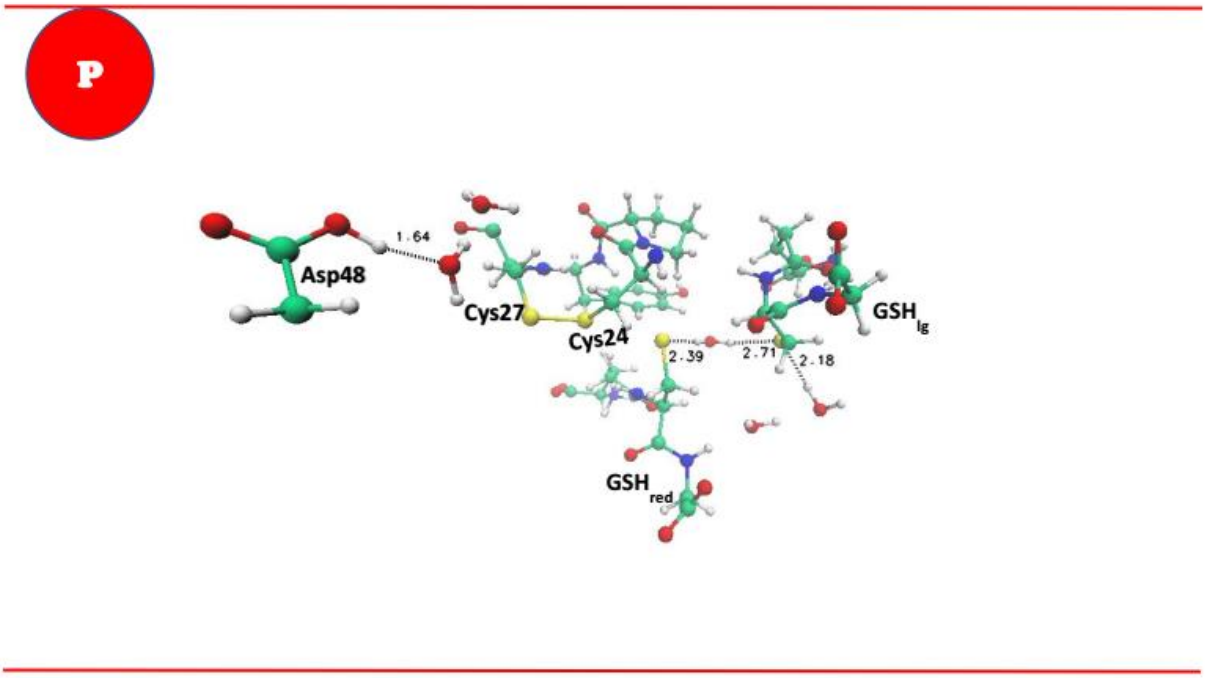



Fig. 24: The reaction states for the intramolecular oxidation of the Cys27 deprotonated mixed disulfide intermediate to form the oxidized Grx and the reduced GSH.

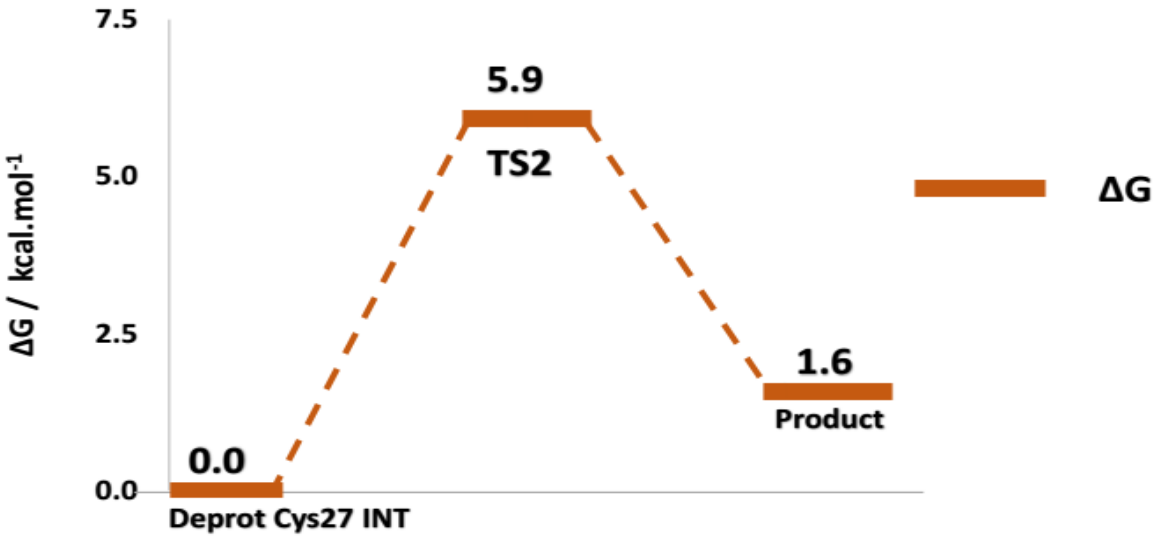



Fig. 25: Thermochemical Profile for the intramolecular oxidation (second thiol-disulfide exchange).

### 3.2.4 Conclusions from the Second Reaction Step

The step involves the cleavage of the Grx-S-S-G mixed disulfide intermediate, that is, intramolecular oxidation through the deprotonated Cys27-thiol to give oxidized Grx ($GrxS_2$) and glutathione (GSH). This reaction step is easy due to the proximity of the buried cysteine (Cys27) to the mixed disulfide intermediate and the stabilization of the products by the environment.

The transition state gave the trisulfide anion ($S_{bur}…S_{nuc}…S_{ctr}$) with bond angle of 169.64° which is close to the linear arrangement necessary for the thiol-disulfide exchange reaction. This TS presents a very low activation free energy barrier of 5.9 $kcal.mol^{-1}$. This barrier is lower than that found for the first thiol-disulfide exchange, implying that the second thiol-disulfide exchange is faster than the first thiol-disulfide exchange.

The product resulted in a Gibbs free energy of 1.6 $kcal.mol^{-1}$ and the two negatively charged reduced glutathiones are stabilized by the nearby water molecules. The activation and the reaction Gibbs energies strongly depend on the relative stabilization of the leaving thiolates by the environment.

# Chapter Four – Conclusions

# 4.0 Conclusions

The purpose of this work was to evaluate and clarify the catalytic mechanism of the reduction of oxidized glutathione (GSSG-disulfide) by glutaredoxin using a multi-level QM/MM computational approach. The ONIOM method as implemented in Gaussian09 package was utilized, described at ONIOM(mPW1N/6-311+G(2d,2p):Amber//mPW1N/6-31G(d):Amber) level of theory.

The mechanism is a two-step reaction: i) biomolecular nucleophilic attack ($S_N2$) of the Cys-24-thiolate to the GSSG-disulfide along the sulfur-sulfur axis to form the mixed disulfide intermediate (Grx-S-S-G) with the release of one molecule of glutathione ($GSH_{lg}$). This is followed by the deprotonation of the buried cysteine (Cys27-thiol) by Asp48 via a bridging water molecule and ii) intramolecular oxidation through the deprotonated Cys-27 thiol to form the oxidized Grx (Grx-$S_2$) and the second reduced glutathione ($GSH_{red}$).

The result of the calculation reveals that the first thiol-disulfide exchange has a higher Gibbs activation barrier (7.8 kcal.mol$^{-1}$) than the second one (5.9 kcal.mol$^{-1}$), indicating that the second reaction is faster than the first reaction. This result predicts the formation of the Grx-S-S-G mixed-disulfide intermediate as the rate limiting step. The Gibbs activation energy of the first thiol-disulfide exchange (7.8 kcal.mol$^{-1}$) is much lower than the experimental result of 17.2 kcal.mol$^{-1}$ calculated from the pseudo-first order rate constant of 1.66 s$^{-1}$. However, this value represents only an upper limit to which the theoretical values have to obey.

The deprotonation of Cys27-thiol gave a Gibbs activation energy of ~25 kcal.mol$^{-1}$. This value is too high to be correct as the proton transfer is expected to be fast with a very low barrier. The deprotonated Cys27-thiol product is more stable than the starting one as it has a lower energy (exergonic). This means that the deprotonation stage is thermodynamically favourable, but not kinetically. Based on this observation, it was assumed that the Cys27-thiol was deprotonated by the Asp48 via the bridging water molecule since the deprotonation stage of this reaction is not a rate determining step, and besides, the mechanism through which the buried cysteine is deprotonated is still yet to be understood.

The second thiol-disulfide exchange reaction occurred more rapidly than the first one with a Gibbs activation energy of 5.9 kcal.mol$^{-1}$. There is (are) no available experimental result(s) in the literature to compare the second thiol-disulfide exchange result with. When this step was studied with the negatively charged $GSH_{lg}$ protonated and when it as not included in the QM layer, no transition state was observed as the energy increased throughout the scan from the reactants to the corresponding products states. The reduced GSH ($GSH_{red}$) and leaving GSH ($GSH_{lg}$) are both negatively charged and are well stabilized by the nearby water molecules. The two GSHs are widely separated (~5 Å). This is because of the repulsive effect between them as they are both negatively charged. The overall reaction has a free energy of 1.6 kcal.mol$^{-1}$, inferring that the reaction is almost thermoneutral.

This study allows for a complete atomistic understanding of the catalytic mechanism of glutaredoxin, and the provided structural details would be useful for future studies on the inhibitory activity of the enzyme.

## References

1. Foloppe, N. and L. Nilsson, The glutaredoxin -C-P-Y-C- Motif: Influence of Peripheral Residues. *Structure*, 2004. **12**: p. 289-300.
2. Eklund, H., Cambillau, C., Sjoberg, B.M., Holmgren, A., Jornvall, H., Hoog, J.O., and Branden, C.I., Conformational and functional similarities between glutaredoxins and thioredoxins. *EMBO J*, 1984. **3**: p. 1443-1449.
3. Martin, J.L., Thioredoxin-a fold for all reasons. *Structure*, 1995. **3**: p. 245-250.
4. Dillet, V., and Brashford, D., Calculations of electrostatic interaction and $pK_{as}$ in the active site of *Escherichia coli* thioredoxin. *Biochemistry*, 1998. **37**: p. 10298-10306.
5. Gan, Z., and Wells, W.W., Identification and reactivity of the catalytic site of pig liver transferase. *J. Biol. Chem*, 1987. **310**: p. 449-470.
6. Holmgren, A., Thioredoxin structure and mechanism: conformational changes on oxidation of the active site sulfhydryls to a disulfide. *Structure*, 1995. **3**: p. 239-243.
7. Kallis, G.B., and Holmgren, A., Differential reactivity of the functional sulfhydryl groups of cysteine-32 and cysteine-35 present in the reduced form of thioredoxin from *Escherichia coli*. *J. Biol. Chem*, 1980. **255**: p. 10261-10265.
8. Kortemme, T., Darby, N.J., and Creighton, T.H., Electrostatic interactions in the active site of the N-terminal thioredoxin-like domain of protein disulfide isomerase. *Biochemistry*, 1996. **35**: p. 14503-14511.
9. Nelson, J.W., and Creighton, T.E., Reactivity and ionization of the active site cysteine residues of DsbA, a protein required for disulfide bond formation in vivo. *Biochemistry*, 1994. **33**: p. 5974-5983.
10. Nordstrand, K., Aslund., F., Meunier, S., Holmgren, A., Otting, G., and Bendt, K.D., Direct NMR observation of the Cys-14 thiol proton of reduced Escherichia coli glutaredoxin-3 supports the presence of an active sites thiol-thiolate hydrogen bonds. *FEBS Lett*., 1999b. **449**: p. 196-200.
11. Yang, Y., and Wells, W.W., Identification and characterization of the functional amino acids at the active center of pig liver thiol-transferase by site-directed mutagenesis. *J. Biol. Chem*., 1991. **266**: p. 12759-12765.
12. Dyson, H.J., Jeng, M.F., Tennant, L.L., Slaby, L., Lindell, M., Cui, D.S., Kuprin, S., and Holmgren, A., Effects of buried charged groups on cysteine thiol ionization and reactivity in *Escherichia coli* thioredoxin: structural and functional characterization of mutants of Asp 26 and Lys 57. *Biochemistry*, 1997. **36**: p. 2622-2636.

13. Jeng, M.F., Holmgren, A., and Dyson, H.J., Proton sharing between cysteine thiols in *Escherichia coli* thioredoxin: implications for the mechanism of protein disulfide reduction. *Biochemistry*, 1995. **34**: p. 10101-10105.
14. Grauschopf, U., Winther, J.R., Korber, P., Zander, T., Dallinger, P., and Bardwell, J.C.A., Why is DsbA such an oxidizing disulfide catalyst? *Cell*, 1995. **83**: p. 947-955.
15. Gilbert, H.F., Molecular and cellular aspects of thiol-disulfide exchange. *Adv. Enzymol. Relat. Areas Mol. Biol.,* 1990. **63**: p. 69-172.
16. Szajewski, R.P., and Whitesides, G.M., Rate constants and equilibrium constants for thiol-disulfide interchange reactions involving oxidized glutathione. *J. Am. Soc*., **1980**. 102: p. 2011-2026.
17. Guddat, L.W., Bardwell, J.C.A., Glockshubber, R., Huber-Wunderlich, M., Zander, T., and Martin, J.L., Structural analysis of three His32 mutants of DsbA: support for an electrostatic role of His32 in DsbA stability. *Protein Sci*., 1997. **6**: p. 1893-1900.
18. Huber-Wunderlich, M., and Glockshuber, R., A single dipeptide sequence modulates the redox properties of a whole enzyme family. *Fold. Des*., 1998. **3**: p. 161-171.
19. Chivers, P.T., Prehoda, K.E., and Raines, R.T., The CXXC motif: a rheostat in the active site. *Biochemistry*, 1997. **36**: p. 4061-4066.
20. Foloppe, N., Segemark, J., Nordstrand, K., Berndt, K.D., and Nilsson, L., Structure, dynamics and electrostatics of the active site of glutaredoxin-3 from *Escherichia coli*: comparison with functionally related proteins. *J. Mol. Biol*., 2001. **310**: p. 449-470.
21. Gane, P.J., Freedman, R.B., and Warwicker, J., A molecular model for the redox potential difference between thioredoxin and DsbA, based on electrostatic calculations. *J. Mol. Biol*., 1995. **249**: p. 376-387.
22. Guddat, L.W., Bardwell, J.C.A., and Martin, J., Crystal structure of the reduced and oxidized DsbA: investigation of domain motion and thiolate stabilization. *Structure*, 1998. **6**: p. 757-767.
23. Weichel., A., Gasdaska, J.R., Powis, G., and Montfort, W.R., Crystal structures of reduced, oxidized, and mutated human thioredoxins: evidence for a regulatory homodimer. *Structure*, 1996. **4**: p. 735-751.
24. Holmgren, A., Hydrogen donor system for *Escherichia coli* ribonucleoside-diphosphate reductase dependent upon glutathione. *Proc. Natl Acad Sci USA*, 1976. **73**: p. 2275-2279.

25. Holmgren, A., Glutathione-dependent synthesis of deoxyribonucleotides. Characterization of the enzymatic mechanism of *Escherichia coli* glutaredoxin. *J. Biol. Chem*., 1979. **254**: p. 3672-3678.
26. Berndt, C., Lillig, C.H., and Holmgren, A., Thioredoxins and glutaredoxins as facilitators of protein folding. *Biochimica et Biophysica Acta,* 2008. **1783**: p. 641-650.
27. Herrero, E., de la Torre-Ruiz, M.A., Monothiol glutaredoxins: a common domain for multiple functions. *Cell. Mol. Life Sci.,* 2005. **6:** p. 1518-1530.
28. Lillig, C.H., Berndt, C., and Holmgren, A., Glutaredoxin systems. *Biochimica et Biophysica Acta,* 2008. **1780**: p. 1304-1317.
29. Lillig, C.H., and Berndt, C., Glutaredoxins in thiol/disulfide exchange. *Antioxidants and Redox Signaling,* 2013. **18(13):** p. 1654-1665.
30. Xing, S., Lauri, A., and Zachgo, S., Redox regulation and flower development: a novel function for glutaredoxin. *Plant Biol. (Stuttg),* 2006. **8:** p. 547-555.
31. Holmgren, A, Thioredoxin and glutaredoxin systems. *J. Biol. Chem*., 1989. **264**: p. 13963-13966.
32. Holmgren, A., Johansson, C., Bendt, C., Lonn, M.E., Hudemann, C., and Lillig, C.H., Thiol redox control via thioredoxin and glutaedoxin systems. *Biochem. Soc. Trans.,* 2005. **33:** p. 1375-1377.
33. Berndt, C., Lillig, C.H., and Holmgren, A., Thiol-based mechanisms of the thioredoxin and glutaredoxin systems: implications for diseases in the cardiovascular system. *Am. J. Physiol, Heart Circ. Physiol.,* 2007. **292**: p. 1227-1236.
34. Collinson, E.J., Wheeler, G.L., Garrido, E.O., Avery, A.M., Avery, S.V., and Grant, C.M., The yeast glutaredoxins are active as glutathione peroxidase. *J. Biol. Chem.,* 2002. **277**: p. 16712-16717.
35. Luikenhuis, S., Perrone, G., Dawes, I.W., and Grant, C.M., The yeast *Saccharomyces cerevisiae* contains two glutaredoxin genes that are required for protection against reactive oxygen species. *Mol. Biol. Cell,* 1998. **9:** p. 1081-1091.
36. Lillig, C.H., Prior, A., Schwenn, J.D., Aslund, F., Ritz, D., Vlamis-Gardikas, A., and Holmgren, A., New thioredoxins and glutaredoxins as electron donors of 3´ phosphoadenylylsulfate reductase. *J. Biol. Chem,* 1999. **274:** p. 7695-7698.
37. Berndt, C., Lillig, C.H., Wollenberg, M., Bill, E., Mansilla, M.C., de Mendoza, D., Seidler, A., and Schwenn, J.D., Characterization and reconstitution of a 4Fe4S adenylyl sulfate

(APS)/phospho-adenylyl sulfate (PAPS) reductase from Bacillus subtilis. *J. Biol. Chem.,* 2004. **279:** p. 7850-7855.

38. Wells, W.W., Xu, D.P., Yang, Y.F., and Rocque, P.A., Mammalian thioltransferase (glutaredoxin) and protein disulfide isomerase have dehydroascorbate reductase activity. *J. Biol. Chem.,* 1990. **265:** p. 15361-15364
39. Takashima, Y., Hirota, K., Nakamura, H., Nakamura, T., Akiyama, K., Cheng, F.S., Maeda, M., and Yodoi, J., Differential expression of glutaredoxin and thioredoxin during monocytic differentiation. *Immunol. Lett.,* 1999. **68:** p. 397-401.
40. Bandyopadhyay, S., Starke, D.W., Mieyal, J.J., and Gronostajski, R.M., Thioltransferase (glutaredoxin) reactivates the DNA-binding activity of oxidation-inactivated nuclear factor I. *J. Biol. Chem.,* 1998. **273:** p. 392-397.
41. Nakamura, T, Ohno, T., Hirota, K., Nishiyama, A., Nakamura, H., Wada, H., and Yodoi, J., Mouse glutaredoxin – cDNA cloning, high level expression in E. coli and its possible implication in redox regulation of the DNA binding activity in transcription factor PEBP2. *Free Radic. Res.,* 1999. **31:** p. 357-365.
42. Hirota, K., Matsui, M., Murata, M., Takashima, Y., Cheng, F.S., Itoh, T., Fukuda, K., and Yodoi, J., Nucleoredoxin, glutaredoxin, and thioredoxin differentially regulate NF-kappaB, AP-1, and CREB activation in HEK293 cells. *Biochem. Biophys. Res. Commun.,* 2000. **274:** p. 177-182.
43. Chrestensen, C.A., Starke, D.W., and Mieyal, J.J., Acute cadmium exposure inactivates thioltransferase (glutaredoxin), inhibits intracellular reduction of protein-glutathionyl-mixed disulfides, and initiates apoptosis. *J. Biol. Chem.,* 2000. **275:** p. 26556-26565.
44. Daily, D., Vlamis-Gardikas, A., Offen, D., Mittlemann, L., Melamed, E., Holmgren, A., and Barzilai, A., Glutaredoxin protects cerebellar granulae neurons from dopamine-induced apoptosis by dual activation of the Ras-phosphoinositide 3-kinase and Jun N-terminal kinase pathway. *J. Biol. Chem.,* 2001. **276:** p. 21618-21626.
45. Daily, D., Vlamis-Gardikas, A., Offen, D., Mittlemann, L., Melamed, E., Holmgren, A., and Barzilai, A., Glutaredoxin protects cerebellar granulae neurons from dopamine-induced apoptosis by activating NF-kappaB via Ref-1. *J. Biol. Chem.,* 2001. **276:** p. 1335-1344.
46. Fernando, M.R., Lechner, J.M., Löfgren, S., Gladyshev, V.N., and Lou, M.F., Mitochondrial thioltransferase (glutaredoxin 2) has GSH-dependent and thioredoxin

reductasedependent peroxidase activities *in vitro* and in lens epithelial cells. *FASEB J.,* 2003. **20:** p. 2645-2647.

47. Starke, D.W., Chock, P.B., and Mieyal, J.J., Glutathione-thiyl radical scavenging and transferase properties of human glutaredoxin (thioltransferase). Potential role in redox signal transduction. *J. Biol. Chem.,* 2003. **278:** p. 14607-14613.
48. Kenchappa, R.S., Diwakar, L., Boyd, M.R., and Ravindranath, V., Thioltransferase (glutaredoxin) mediates recovery of motor neurons from excitotoxic mitochondrial injury. *J. Neurosci.,* 2002. **22:** p. 8402-8410.
49. Xing, K., and Lou, M.F., The possible physiological function of thioltransferase in cells. *FASEB J.,* 2003. **17:** p. 2088-2090.
50. Wang. Y., Qiao, M., Mieyal, J.J., Asmis, L.M., and Asmis, R., Molecular mechanism of glutathione-mediated protection from oxidized low-density lipoprotein-induced cell injury in human macrophages: role of glutathione reductase and glutaredoxin. *Free Radic Biol Med.,* 2006. **41:** p. 775-785.
51. Davis, D.A., Newcomb, F.M., Starke, D.W., Ott, D.E., Mieyal, J.J., and Yarchoan R., Thioltransferase (glutaredoxin) is detected within HIV-1 and can regulate the activity of glutathionylated HIV-1 protease *in vitro*. *J. Biol. Chem.,* 1997. **272:** p. 25935-25940.
52. Rodriguez-Manzaneque, M.T., Tamarit, J., Belli, G., Ros, J., Herrero, E., Grx5 is a mitochondrial glutaredoxin required for the activity of iron/sulfur enzymes. *Mol. Biol. Cell,* 2002. **13:** p. 1109-1121.
53. Sun, C., Berardi, M.J., and Bushweller, J.H., The NMR solution of human glutaredoxin in the fully reduced form. *J. Mol. Biol.,* 1998. **280:** p. 687-701.
54. Berardi, M.J., Pendred, C.L., and Bushweller, J.H., Preparation, characterization and complete heteronuclear NMR resonance assignments of the glutaredoxin (C14S)-ribonucleotide reductase B1 737-761 (C754S) mixed disulfide. *Biochemistry,* 1998. **37:** p. 5849-5857.
55. Klintrot, I.M., Hoog, J.O., Jornvall, H., Holmgren, A., and Luthman, M., The primary structure of calf thymus glutaredoxin. Homology with the corresponding *Escherichia coli* protein but elongation at both ends with an additional half-cysteine/cystine pair. *Eur. J. Biochem,* 1984. **144:** p. 417-423.
56. Fernandes, P.A., and Ramos, M.J., Theoretical insights into the mechanism for thiol/disulfide exchange. *Chem. Eur. J.,* 2004. **10:** p. 257-266.

57. Gallogly, M.M., Starke, D.W., and Mieyal, J.J., Mechanistic and kinetic details of catalysis of thiol-disulfide exchange by glutaredoxins and potential mechanisms of regulation. *Antioxidants and Redox Signaling,* 2009. **11(5):** p. 1059-1081.
58. Gravina, S.A., and Mieyal, J.J., Thioltransferase is a specific glutathionyl mixed disulfide oxidoreductase. *Biochemistry,* 1993. **32:** p. 3368-3376.
59. Carvalho, A.P., Fernandes, P.A., and Ramos, M.J., Similarities and differences in the thioredoxin superfamily. *Prog. Biophys. Mol. Biol.,* 2006. **91**: p. 229-248.
60. Adachi, T., Weisbrod, R.M., Pimentel, D.R., Ying, J, Sharov, V.S., Schöneich, C., and Cohen, R.A., S-Glutathiolation by peroxynitrite activates SERCA during arterial relaxation by nitric oxide. *Nat. Med.,* 2004. **10**: p. 1200-1207.
61. Aslund, F., Berndt, K.D., and Holmgren, A., Redox potentials of glutaredoxins and other thioldisulfide oxidoreductases of the thioredoxin superfamily determined by direct protein-protein redox equilibria. *J. Biol. Chem.,* 1997. **272**: p. 30780–30786.
62. Srinivasan, U., Mieyal, P.A., and Mieyal, J.J., pH profiles indicative of rate limiting nucleophilic displacement in thioltransferase catalysis. *Biochemistry*, 1997. **36**: p. 3199-3206.
63. Gallogly, M.M., Starke, D.W., Leonberg, A.K., Ospina, S.M., and Mieyal, J.J., Kinetic and mechanistic characterization and versatile catalytic properties of mammalian glutaredoxin 2: implications for intracellular roles. *Biochemistry,* 2008. **47**: p. 11144–11157.
64. Carvalho, A.P., Fernandes, P.A., Swart, M., Van Stralen, J.N.P., Bickelhaupt, F.M., and Ramos, M.J., Role of the variable active site residues in the function of thioredoxin family oxidoreductases. *J. Comp. Chem.,* 2009. **30(5)**: p. 710-724.
65. Menchise, V., Corbier, C., Didierjean, C., Saviano, M., Benedetti, E., Jacquot, J.P., and Aubry, A., Crystal structure of the wild-type and D30A mutant thioredoxin h of Chlamydomonas reinhardtii and implications for the catalytic mechanism. *Biochem. J.,* 2001. **359**: p. 65-75.
66. Rickard, G.A., Berges, J., Houee-Levin, C., and Rauk, A., Ab Initio and QM/MM study of electron addition on the disulfide bond in thioredoxin. *J. Phys. Chem. B,* 2008. **112(18):** p. 5774-5787.
67. Carvalho, A.P., Swart, M., Van Stralen, J.N., Fernandes, P.A., Ramos, M.J., and Bickelhaupt, F.M., Mechanism of thioredoxin-catalysed disulfide reduction. Activation of

the buried thiol and role of the variable active-site residues. *J. Phys. Chem. B,* 2008. **112(18)**: p. 2511-2523.

68. Neves, R.P.P., Fernandes, P.A., and Ramos, M.J., Mechanistic insights on the reduction of glutathione disulfide by protein disulfide isomerase. *PNAS,* 2017. p. E4724-E4733.
69. Pappas, J.A., Theoretical studies of the reactions of the sulfur-sulfur bond. General heterolytic mechanisms. *J. Am. Chem. Soc.,* 1977. **99:** p. 2926–2930.
70. Hol, W.G., The role of the alpha-helix dipole in protein function and structure. *Prog. Biophys. Mol. Biol.,* 1985. **45:** p. 149-195.
71. Katti, S.K., LeMaster, D.M., and Eklund, H., Crystal structure of thioredoxin from *Escherichia coli* at 1.68Å resolution. *J. Mol. Biol.,* 1990. **212:** p. 167-184.
72. Gane, P.J., Freedman, R.B., and Warwicker, J., A molecular model for the redox potential difference between thioredoxin and DsbA based on electrostatic calculations. *J. Mol. Biol.,* 1995. **249:** p. 376-387.
73. Kortemme, T., and Creighton, T., Ionisation of cysteine residues at the termini of model alpha-helical peptides. Relevance to unusual thiol pKa values in proteins of the thioredoxin family. *J. Mol. Biol.,* 1995. **253:** p. 799-821.
74. Jeng, M., Holmgren, A., and Dyson, H.J., Proton sharing between cysteine thiols in *Escherichia coli* thioredoxin: implication for the mechanism of protein disulfide reduction. *Biochemistry,* 1995. **34:** p. 10101-10105.
75. Forman-Kay, J.D., Clore, G.M., and Gronenborn, A.M., Relationship between electrostatics and redox function in human thioredoxin: characterization of pH titration shifts using two-dimensional homo- and heteronuclear NMR. *Biochemistry,* 1992. **31:** p. 3442-3452.
76. Eklund, H., Ingelman, M., Soderberg, B.O., Uhlin, T., Nordlund, P., Nikkola, M., Sonnerstam, U., Joelson, T., and Petratos, K., Structure of oxidized bacteriophage T4 glutaredoxin (thioredoxin). *J. Mol. Biol.,* 1992. **228:** p. 596-618.
77. Sun, C., Holmgren, A., and Bushweller, J.H., Complete $^{1}H$, $^{13}C$ and $^{15}N$ NMR resonance assignments and secondary structure of human glutaredoxin in the fully reduced form. *Protein Sci.,* 1997. **6:** p. 383-390.
78. Lundstrom-Ljung, J., Krause, G., and Holmgren, A., A Pro to His mutation in active site of thioredoxin increases its disulfide-isomerase activity 10-fold. *J. Biol. Chem.,* 1992. **267:** p. 9047-9052.

79. Jao, S., Ospina, S.M.E., Berdis, J., Starke, D.W., Post, C.B., and Mieyal, J.J., Computational and mutational analysis of human glutaredoxin (thioltransferase): Probing the molecular basis of the low pKa of Cysteine 22 and its role in catalysis. *Biochemistry,* 2006. **45:** p. 4785-4796.

80. Zaffagnini, M., *et al.,* Glutaredoxin s12: unique properties for redox signaling. *Antioxid. Redox Signal,* 2012. **16:** p. 17-32.

81. Djuika, C.F., et al., Plasmodium falciparum antioxidant protein as a model enzyme for a special class of glutaredoxin/glutathione-dependent peroxiredoxins. *Biochim. Biophys. Acta,* 2013. **1830:** p. 4073-4090.

82. Toppo, S., Flohe, L., Ursini, F., Vanin, S., and Maiorino, M., Catalytic mechanisms and specificities of glutathione peroxidases: variations of a basic scheme. *Biochim. Biophys. Acta,* 2009. **1790:** p. 1486-1500.

83. Ramos, M.J., and Fernandes, P.A., Computational enzymatic catalysis. *Acc. Chem. Resear.,* 2008. **41**(6**)**: p. 689-698.

84. Lonsdale, R., Harvey, J.N., and Mulholland, J. A practical guide to modelling enzyme-catalysed reactions. *Chem. Soc. Rev.,* 2012. **41:** p. 3025-3038.

85. Thiel, W. Semiempirical quantum-chemical methods. *WIREs Comput Mol Sci.,* 2013, 00: 1-10 doi: 10.1002/wcms.1161.

86. Young, D. Computational Chemistry: A practical guide for applying techniques to real world problems. *Wiley-Interscience,* 2001, New York.

87. Atkins, P.W. and Friedman, R.S. Molecular Quantum Mechanics,3rd Ed., 1997, Oxford.

88. Foresman, J.B., and Frisch, A. Exploring chemistry with electronic structure methods, 2nd Ed., 1996, *Gaussian, Inc.,* Pittsburgh: PA.

89. Leach, A.R. Molecular modeling principles and applications. Addison Wesley Longman.

90. Thiel, W. Semiempirical quantum-chemical methods in computational chemistry. *Theory and Applications of Computational Chemistry,* 2005. p. 577-580.

91. Bowman, A.L., Grant, I.M., and Mulholland, A.J. QM/MM simulations predict a covalent intermediate in the hen egg white lysozyme reaction with its natural substrate. *Chem. Commun.,* 2008. **37:** p. 4425-4427.

92. Dewar, M.J.S., Zoebisch, E., Healy, E.F., and Stewart, J.J.P. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model, *J. Am. Chem. Soc.,* 1985. **107:** p. 3902-3909.

93. Stewart, J.J.P. Optimization of parameters for semiempirical methods I. Methods. *J. Comput Chem.,* 1989. **10:** p. 209-220.
94. Stewart, J.J.P. Optimization of parameters for semiempirical methods II. Applications. *J. Comput Chem.,* 1989. **10:** p. 221-264.
95. Elstner, M., Frauenheim, T., and Suhai, S. THEOCHEM. 2003. **632:** p. 29-41.
96. Seifert, G., and Joswig, J.O. Density functional tight-binding – an approximate DFT method. *WIREs Comp. Mol. Sci.,* 2012. **2:** p. 456-465.
97. Elstner, M., Gaus, M., Cui, Q. Density functional tight binding: applications to organic and biological molecules. *WIREs Comp. Mol. Sci.,* 2013, doi: 10.1002/wcms.1156.
98. Otte, N., Scholten, M., and Thiel, W. Looking at self-consistent-charge density functional tight binding from a semiempirical perspective. *J. Phys. Chem. A.* 2007. **111:** p. 5751-5755.
99. Born, M., and Huang, K. Dynamical theory of crystal lattices, 1954, Oxford University Press, Oxford.
100. Foulkes, W.M.C., Mitas, L., Needs, R.J., and Rajagopal, G. 2001. Reviews of Modern Physics 73, 33.
101. Cramer, C.J. Essential of computational chemistry, theories and models, second Edition. John Wiley & Sons Ltd, West Sussex 2004.
102. Johnson, B.G., Gill, P.M.W., and Pople, J.A. The performance of a family of density functional methods. *J. Chem. Phys.,* 1993. **98** (7): p. 5612-5626.
103. Frank Jensen. Introduction to computational chemistry. John Wiley & Sons Ltd, West Sussex 2007.
104. Sousa, S.F., Fernandes, P.A., and Ramos, M.J. General performance of density functionals. *J. Phys. Chem. A,* 2007. **111:** p. 10439-10452.
105. Kohn, W., and Sham, L.J. Self-consistent equations including exchange and correlation effects. *Phys. Rev. A,* 1965. **140:** p. 1133.
106. Perdew, J.P. Climbing the ladder of density functional approximation. MRS Bulletin, 2013. **38**(09): p. 743-750.
107. Cohen, A.j., Mori-Sanchez, P., and Yang, W. Challenges for density functional theory. *Chem. Rev.,* 2012. **112**(1): p. 289-320.
108. Perdew, J.P. Jacob's ladder of density functional approximations for the exchange-correlation energy. 2001. **577:** p. 1-20.

109. Langreth, D.C., and Mehl, M.J. Beyond the local density approximation in calculations of ground state electronic properties. *Phys. Rev. B.* 1983. **28**(4): p. 1809.

110. Becke, A.D. Density functional exchange energy approximation with correct asymptotic behavior. *Phys. Rev. A.* 1988. **38**: p. 3098.

111. Perdew, J.P., Chevary, J.A., Vosko, S.H., Jackson, K.A., Pederson, M.R., Singh, D.J., and Fiolhais, C. Atoms, molecules, solids, and surfaces: applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B.* 1992. **46**: p. 6671.

112. Becke, A.D. Density functional thermochemistry. I. the effect of the exchange-only gradient correction. *J. Chem. Phys.,* 1992. **96**: p. 2155.

113. Proynov, E.I., Ruiz, E., Vela, A., and Salahub, D.R. Determining and extending the domain of exchange and correlation functionals. *Int. J. Quantum Chem.,* 1995. **29:** p. 61.

114. Tao, J., Perdew, J.P., Staroverov, V.N., and Scuseria, G.E. Climbing the density functional ladder: nonempirical meta-generalized gradient approximation designed for molecules and solids. *J. Chem. Phys.,* 2005. 122, Art. No. 114102.

115. Patton, D.C., and Pederson, M.R. Application of the generalized gradient approximation to rare-gas dimers. *Phys. Rev. A.* 1997. **56:** p. 2495.

116. Dunning, T.H. *J. Phys. Chem.,* 1989. **90:** p. 1007.

117. Kostal, J. Computational chemistry in predictive toxicology: status quo et quo vadis? *Adv. Mol. Tox.,* 2017. **10:** p. 139-183.

118. Wathen, B., Pratt, D.A., and Jia, Z. Hyperconjugation contributes to the biomodal distribution of glycine conformations observed in protein three-dimensional structures. *ChemBioChem.,* 2011. **12**(11): p. 1674-1677.

119. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S, and Karplus, M. CHARMM – a program for macromolecular energy minimization, and dynamics simulations. *J. Comp. Chem.,* 1983. **4:** p. 187-217.

120. MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J. *et al*. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B,* 1998. **102:** p. 3586-3616.

121. MacKerell, A.D., Feig, M., and Brooks, C.L. Extending the treatment of the backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics. *J. Comput. Chem.,* 2004. **25:** p. 1400-1415.

122. Kaminski, G., and Jorgensen, W.L. Performance of the AMBER94, MMFF94, and OPLS-AA force fields for modeling organic liquids. *J. Phys. Chem.,* 1996. **100:** p. 18010-18013.

123. Vanommeslaeghe, K., Hatcher, E., Acharya, C. *et al.,* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.,* 2010. **31:** p. 671-690.

124. Maxwell, D.S., Tirado-Rives, J., and Jorgensen, W.L. A comprehensive study of the rotational energy profiles of organic system by ab initio MO theory, forming a basis for peptide torsional parameters. *J. Comput. Chem.,* 1995. **16:** 984-1010.

125. Weiner, S.J., Kollman, P.A., Case, D.A. et al., A new force field for molecular mechanical simulation of nuclei-acids and proteins. *J. Am. Chem. Soc.,* 1984, **106:** 765-784.

126. Zhang, Y., Sagui, C. The gp41(659-671) HIV-1 antibody epitope: a structurally challenging small peptide. *J. Phys. Chem. B.* 2014. **118:** p. 69-80.

127. Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.,* 2004. **25:** p. 1157-1174.

128. Gao, J., Ma, S., Major, D.T., Nam, K., Pu, J., and Truhlar, D.G. Mechanisms and free energies of enzymatic reactions. *Chem. Rev.,* 2006. **106:** p. 3188-3209.

129. Rosta, E., Klahn, M., Warshel, A. Towards accurate ab initio QM/MM calculations of free-energy profiles of enzymatic reactions. *J. Phys. Chem B,* 2006. **110:** p. 2934-2941.

130. Warshel, A., Levitt, M. Theoretical studies of enzymatic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of Lysozyme. J. Mol. Bio. 1976. 103: p. 227-249.

131. Gao, J.L., Truhlar, D.G. Quantum mechanical methods for enzyme kinetics*. Annu. Rev. Phys. Chem.,* 2002. **53:** p. 467-505.

132. Groenhof, G. Introduction to QM/MM Simulations. *Methods Mol. Biol.,* 2013. **924:** p. 43-66.

133. Oliveira, E.F., Cerqueira, N.M.F.S.A., Fernandes, P.A., and Ramos, M.J. Mechanism of formation of the internal Aldimine in pyridoxal 5'-phosphate dependent enzymes. *J. Am. Chem. Soc.,* 2011. **133**(39): p. 15496-15505.

134. Gesto, D.S., Cerqueira, N.M.F.S.A., Fernandes, P.A., and Ramos, M.J. Unravelling the enigmatic mechanism of L-Asparaginase II with QM/QM calculations. *J. Am. Chem. Soc.,* 2013. **135**(19): p. 7146-7158.

135. Andrejic, M.A., and Mata, R. Local hybrid QM/QM Calculations of reaction pathways in metallobiosites. *J. Chem. Theory Comput.,* 2014. **10**(12): p. 5397-5404.

136. Cerqueira, N.S.., Fernandes, P.A., and Ramos, M.J. Protocol for computational enzymatic reactivity based on geometry optimization. *ChemPhysChem* 10.1002/cphc.201700339.

137. Senn, H.M., and Thiel, W. QM/MM method for biomolecular systems. *Angew. Chem. Int. Ed.* 2009. **48:** p. 1198-1229.

138. Karelina, M., and Kulik, H.J. Systematic quantum mechanical region determination in QM/MM simulation. J. Chem. Theory Comput., 2017, doi: 10.1021/acs.jcts.6b01049.

139. Sousa, S.F., Ribeiro, A.J.M., Neves, R.P.P., Bras, N.F., Cerqueira, N.M.F.S.A., P.A., Fernandes, and Ramos, M.J. Application of quantum mechanics/molecular mechanics methods in the study of enzymatic reaction mechanisms. *WIREs Comput. Mol. Sci* 2017, 7:e1281. doi: 10.1002/wcms.1281.

140. Chung, L.W., Hirao, H., Li, X., Morokuma, K. The ONIOM Method: its foundation and applications to metalloenzymes and photobiology. *WIREs Comput. Mol. Sci.,* 2012. **2:** p. 327-350.

141. Maseras, F., Morokuma, K. IMOMM – a new integrated ab initio plus molecular mechanic geometry optimization scheme of equilibrium structures and transition states. *J. Comput. Chem.,* 1995. **16:** p. 1170-1179.

142. Humbel, S., Sieber, S., Morokuma, K. The IMOMO method: integration of different levels of molecular orbital approximation for geometry optimization of large systems: test for n-butane conformation and SN2 reaction: RCl+Cl. *J. Chem. Phys.,* 1996. **105:** p. 1959-1967.

143. Svensson, M., Humbel, S., Froese, R.D.J., Matsubara, T., Sieber, S., and Morokuma, K. ONIOM: a multilayered integrated MO+MM method for geometry optimizations and single point energy predictions. A test for Dield-Alder reactions and Pt(P(t-Bu)(3))(2)+H-2 oxidative addition. *J. Phys. Chem.,* 1996. **100:** p. 19357-19363.

144. Kamaromi, I., and Muszbek, L. In frontiers of multifunctional nanosystems (NATO Sci. Ser. II, Vol. 57) (Eds.: E.V. Buzaneva, P. Scharff), Kluwer, Dordrecht, 2002. P. 17-28.

145. Vreven, T., K.S., Byun, Komaromi, I., *et al.,* Combining quantum mechanics methods with molecular mechanic methods in ONIOM. *J. Chem. Theory Comput.,* 2006. 2: p. 815-826.

146. Chung, L.W., Sameera, W.M.C., Ramozzi., R., Page, A.J., et al., The ONIOM Method and its application. Chem. Rev. 2014, doi: 10.1021/cr5004419.

147. Lin, H., and Truhlar, D.G. QM/MM: what have we learned, where are we, and where do we go from here? *Theor. Chem. Acc.,* 2007. **117:** p. 185-199.
148. Assfeld, X., and Rivail., J. Quantum chemical computations on parts of large molecules: the ab initio local self-consistent field method. *Chem. Phys. Lett.,* 1996. **263:** p. 100-106.
149. Thery, V., Rinaldi, D., Rivail, J.L., Maigret, B., and Ferenczy, C.G. Quantum mechanical computations on very large molecular systems – the local self-consistent field method. *J. Comput. Chem.,* 1994. **15:** p. 269-282.
150. Gao, J., Patricia, A., Alhambra, C., and Field, M.J. A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations. *J. Phys. Chem. A.* 1998. **102**(24): p. 4714-4721.
151. Bayly, C., Cieplak, P., Cornell, P., and Kollman, P. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges – the RESP model. *J. Phys. Chem.*, 1993. **97:** p. 10269-10280.
152. Besler, B., Merz, K., and Kollman, P. Atomic charges derived from semiempirical methods. *J. Comput. Chem.,* 1990. **11:** p. 431-439.
153. Bakowies, D., and Thiel, W. Hybrid models for combined quantum mechanical and molecular mechanical approaches. *J. Phys. Chem.,* 1996. **100:** p. 10580-10594.
154. Javier, L.F., Dehez, F., Chipot, C., and Orozco, M. Polarization effects in molecular interactions. *WIREs Comput Mol. Sci.,* 2011. **1:** p. 844-854.
155. Ponder, J.W., and Case, D.A. Force fields for protein simulations. *Advances in protein chemistry,* 2003. **66:** p. 27-85
156. Kaminski, G.A., Stern, H.A., Berne, B.J., and Friesner, R.A. Development of an accurate and robust polarizable molecular mechanics force field from ab initio quantum chemistry. *J. Phys. Chem. A,* 2004. **108:** 621-627.
157. Lopes, P.E.M., Huang, J., Shim, J., Luo, Y., Li, H., Roux, B., MacKerell, A.D. Polarizable force field for peptides and proteins based on classical drude oscillator. *J. Chem. Theory Comput.,* 2013. **9:** p. 5430-5449.
158. Boulanger, E., and Thiel, W. Towards QM/MM simulation of enzymatic reactions with the drude oscillator polarizable force field. *J. Chem. Theory Comput.,* 2014. **10:** p. 1795-1809.
159. Cerqueira, N.M.F.S.A., Fernandes, P.A., and Ramos, M.J. Enzyme ribonucleotide reductase: unraveling an enigmatic paradigm of enzyme inhibition by furanone derivatives. *J. Phys. Chem. B,* 2006. **110**(42): p. 21272-21281.

160. Mulholland, A. Computational enzymology: modelling the mechanisms of biological catalysts. *Biochem. Soc. Trans.,* 2008. **36:** p. 22-26.

161. Parker, A.R., Petluru, N.P., Nienaber, V.L. et al., Cysteine specific targeting of the functionally distinct peroxiredoxin and glutaredoxin proteins by the investigational disulfide BNP7787. *Molecules*, 2015. **20**: p. 4928-4950.

162. Lee, E.H., Kim, H.Y., and Hwang, K.Y., The GSH- and GSSG-bound structures of glutaredoxin from Clostridium oremlandii. *Arch Biochem Biophys,* 2014. **565:** p. 20-25.

163. Case, D.A., Darden, T.A., Cheatham, T.E., et al., AMBER 12; University of California, San Francisco, CA, 2012.

164. Gordon, J.C., Myers, J.B., Folta, T., Shoja, V., Heath, L.S., and Onufriev, A., H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res,* 2005. **33:** p. W368-W371.

165. Calixto, A.R., Bras, N.F., Fernandes, P.A., Ramos, M.J., Reaction mechanism of human renin studied by QM/MM calculations. *ACS Catal*. 2014. **4:** p. 3869-3876.

166. Liu, F., Zhao, Y.L., Wang, X, et al., Elucidation of enzymatic mechanism of Phenazine Biosynthetic protein using QM/MM and MD simulations. *Plos One* 2015, 10, e0139081.

167. Neves, R.P.P., Fernandes, P.A., Varandas, A.J.C., Ramos, M.J., Benchmarking of density functionals for the accurate description of thiol-disulfide. *J Chem Theory Comput* 2014. **10:** p. 4842-4856.

168. Ochterski, J.W. Thermochemistry in Gaussian, 2000, p. 19.

169. Deponte, M. Glutathione catalysis and the reaction mechanisms of glutathione-dependent enzymes. *Bioch et Biophys,* 2013. **1830:** p. 3217-3266.

170. Boyd, R.J., Perkyns, J.S., Ramani, R., Conformations of simple disulfides and L-cysteine, *Can J Chem* 1983. **61:** p. 1082-1085.

171. Rabenstein, D.L., and Millis, K.K. Nuclear magnetic resonance study of the thioltransferase catalysed glutathione/glutathione disulfide interchange reaction. *Biochim et Biophys*, 1995. **1249:** p. 29-36.